Ancestral lineages and limit theorems for branching Markov chains

Vincent Bansaye *

November 10, 2014

Abstract

We consider a branching model in discrete time for structured population in varying environment. Each individual has a trait, which belongs to some general state space and both the reproduction law and the trait inherited by the offsprings may depend on the trait of the mother and the environment. We study the long time behavior of the population and the ancestral lineage of typical individuals under general assumptions. We focus on the growth rate and the trait distribution among the population for large time and provide some estimations of the local densities. A key role is played by well chosen (possibly non-homogeneous) Markov chains. It relies in particular on an extension of many-to-one formula [G07, BDMT11] and the analysis of the genealogy, in the vein of the spine decomposition of [LPP95, KLPP97, GB03]. The applications use the spectral gap of the mean operator, the Harris ergodicity or the large deviations of this auxiliary Markov chain.

Key words. Branching processes, Markov chains, Varying environment, Genealogies.

A.M.S. Classification. 60J80, 60J05, 60F05, 60F10

1 Introduction

We are interested in a branching Markov chain, which means a multitype branching process whose number of types may be infinite. The environment may evolve (randomly) but when the environment is given, each individual evolves independently and the quenched branching property hold.

Let (E, T) be a pair consisting of a set E of environments and an invertible map T on E. One can keep in mind the case when the environment is $\mathbf{e} = (e_i : i \in \mathbb{Z})$ and $T\mathbf{e} = (e_{i+1} : i \in \mathbb{Z})$.

*CMAP, Ecole Polytechnique, CNRS, route de Saclay, 91128 Palaiseau Cedex-France; E-mail: vincent.bansaye@polytechnique.edu

Let $(\mathcal{X}, \mathcal{B}_{\mathcal{X}})$ be a measurable space. It gives the state space of the branching Markov chain. The example $\mathcal{X} \subset \mathbb{R}^d$ endowed with Borelian sets will be relevant for the applications.
For each $k \in \mathbb{N}$ and $\mathbf{e} \in E$, let $P^{(k)}(., \mathbf{e}, .)$ be a function from $\mathcal{X} \times \mathcal{B}_{\mathcal{X}^k}$ into $[0, 1]$ which satisfies
a) For each $x \in \mathcal{X}$, $P^{(k)}(x, \mathbf{e}, .)$ is a probability measure on $(\mathcal{X}, \mathcal{B}_{\mathcal{X}^k})$.
b) For each $A \in \mathcal{B}_{\mathcal{X}^k}$, $P^{(k)}(., \mathbf{e}, A)$ is a $\mathcal{B}_{\mathcal{X}}$ measurable function.

In the whole paper, we use the classical following notations for discrete trees. Each individual in the population is an element of $\{\varnothing\} \cup \bigcup_{n \geq 1} (\mathbb{N}^*)^n$ and is denoted by $u = u_1 u_2 ... u_n$ with $u_i \in \mathbb{N}^* = \{1, 2, \dots\}$. We denote by $|u| = n$ the generation of the individual u, by $N(u)$ the number of offsprings of the individual u and by $X(u) \in \mathcal{X}$ the trait (or position) of the individual u.

For any generation, each individual with trait $x \in \mathcal{X}$ which lives in environment $\mathbf{e} \in E$ gives birth independently to a random number of offsprings, whose law both depend on x and $\mathbf{e}$. This number of offsprings is distributed as a r.v. $N(x, \mathbf{e})$ whose mean is denoted by

$$m(x, \mathbf{e}) = \mathbb{E}(N(x, \mathbf{e})).$$

In the whole paper, we assume that $m(x, \mathbf{e}) > 0$ for each $x \in \mathcal{X}, \mathbf{e} \in E$. For the models mentioned here, one can keep in mind $\mathbf{e} = (e_i : i \in \mathbb{Z})$ and $N(x, \mathbf{e})$ depends only on x and e_0.
If the environment is $\mathbf{e}$, we denote by $\mathbb{P}_{\mathbf{e}}$ the associated probability. The distribution of the traits of the offsprings of the individual u living in generation n ($|u| = n$) is given by

$$\begin{aligned}\mathbb{P}_{\mathbf{e}}(X_{u1} \in dx_1, \dots, X_{uk} \in dx_k \mid (X(v) : |v| \leq n), N(u) = k) \\ = P^{(k)}(X(u), T^n \mathbf{e}, dx_1 \dots, dx_k).\end{aligned}$$

Thus, one individual with trait x living in environment $\mathbf{e}$ gives birth to a set of individuals $(X_1, \cdots, X_{N(x,\mathbf{e})})$ whose trait are specified by the transition kernels$(P^{(k)}(\mathbf{e}, x, .) : k \in \mathbb{N}, \mathbf{e} \in E)$.
The process X is a multitype branching process in varying environment where the types take value in $\mathcal{X}$. They have been largely studied for finite number of types, whereas much less is known or understood in the infinite case, but some results due to Seneta, Vere Jones, Moy, Kesten for countable many types.
The case of branching random walk has also attracted lots of attention from the pioneering works of Biggins. Then $\mathcal{X} = \mathbb{R}^d$ and the transitions $P^{(k)}$ are invariant by translation, i.e. $P^{(k)}(x, \mathbf{e}, x + dx_1 \cdots, x + dx_k)$ does not depend on $x \in \mathcal{X}$. Recently, fine results have been obtained about the extremal individuals and their genealogy for such models, see e.g. [HS09, AS10]. and branching random walk in random environment have been investigated. In particular the recurrence property [**?**, CP07a], the survival and the growth rate [GMPV10, CP07b, CY11], central limit theorems [Y08, N11] and large deviations results [HL11] have been obtained.
As far as I see, the classical methods relying on the spectral theory and the martingale arguments are not easily adaptable to the general framework we consider. We are motivated by applications to models for biology and ecology such as cell division models for cellular aging [G07] or parasite infection [B08] and reproduction-dispersion models in non-homogeneous environment [BL12]. Thus, we are also inspired by the utilization of

auxiliary Markov chains, branching decomposition and L^2 computations, in the vein of the works of Athreya and Khang [AK98a, AK98b] and Guyon [G07]. The applications and references will be given along the paper.

We are interested in the evolution of the measure associated to the traits of the individuals:

$$Z_n := \sum_{|u|=n} \delta_{X(u)}$$

and more specifically by $Z_n(A_n) = \#\{u : |u| = n,\ X(u) \in A_n\}$. We also define

$$Z_n(f) = \sum_{|u|=n} f(X(u)), \quad f_n.Z_n = \sum_{|u|=n} \delta_{f_n(X(u))}.$$

First, we want to know if the process may survive globally and how it would then grow. Thus, Section 2 yields an expression of the mean growth rate of the population relying on the dynamic of the trait and the offspring laws, see [BL12] for motivations for metapopulations. Then (Section 3), we study the repartition of the population and focus on the asymptotic behavior of the proportions of individuals whose trait belongs to A, i.e. $Z_n(\mathcal{X})/Z_n(A)$. It is inspired by [AK98a, G07, BH13] and extends the law of large numbers to both varying environment and trait dependent reproduction. It is achieved by introducing a non-homogeneous auxiliary Markov chain (Section 3.1). We add that we take into account some possible renormalization of the traits via a function f_n to cover non recurrent positive cases. Finally, in Section 4, we provide some asymptotic results about $Z_n(A_n)$, outside the range of law of large numbers. It relies on the large deviations of the auxiliary process and the trajectory associated with. As an application we can derive the position of the extremal particles in some monotone models motivated by biology, where new behaviors appear.

We end up the introduction with recalling some classical notations. If $u = u_1 \cdots u_n$ and $v = v_1 \cdots v_m$, then $uv = u_1 \cdots u_n v_1 \cdots v_m$. For two different individuals u, v of a tree, write $u < v$ if u is an ancestor of v, and denote by $u \wedge v$ the nearest common ancestor of u and v in the means that $|w| \leq |u \wedge v|$ if $w < u$ and $w < v$. Moreover, we note $\mathcal{M}(X)$, resp. $\mathcal{M}_f(X)$ and $\mathcal{M}_1(X)$ the set of (non-negative) measures, resp. finite measure and probability distribution of a measurable space X.

2 Growth rate of the population

We denote by $\rho_{\mathbf{e}} = \lim_{n\to\infty} n^{-1} \log \mathbb{E}_{\mathbf{e}}(Z_n(\mathcal{X}))$ the growth rate of the population in the environment $\mathbf{e}$, when it exists.
We are giving an expression of this growth rate in terms of a Markov chain associated with a random lineage. Its transition kernel is defined by

$$P(x, \mathbf{e}, dy) := \frac{1}{m(x, \mathbf{e})} \sum_{k\geq 1} \mathbb{P}(N(x, \mathbf{e}) = k) \sum_{i=1}^{k} P^{(k)}(x, \mathbf{e}, \mathcal{X}^{i-1} dy \mathcal{X}^{k-i})$$

so that the auxiliary Markov chain X is given by

$$\mathbb{P}_{\mathbf{e}}(X_{n+1} \in dy \,|X_n = x) = P(x, T^n\mathbf{e}, dy).$$

It means that we follow a linage by choosing uniformly at random one of the offsprings at each generation, biased by the number of children.

We assume now that $\mathcal{X}$ is a locally compact polish space endowed with a complete metric and its Borel σ field. Moreover E is a Polish Space and T is an homomorphism. In the rest of the paper, we endow $\mathcal{M}_1(\mathcal{X}\times E)$ with the weak topology, where $\mathcal{M}_1(\mathcal{X}\times E)$ is the space of probabilities on $\mathcal{X}\times E$. It is the smallest topology such that $\mu \in \mathcal{M}_1(\mathcal{X}\times E) \to \int_{\mathcal{X}\times E} f(z)\mu(dz)$ is continuous as soon as f is continuous and bounded.

Definition 1. *We say that X satisfies a Large Deviation Principle (LDP) with good rate function $I_{\mathbf{e}}$ in environment $\mathbf{e}$ when there exists a lower semi-continuous function $I : \mathcal{X} \times E \to \mathbb{R}$ with compact level subsets*[1] *for the weak topology such that*

$$L_n^{\mathbf{e}} = \frac{1}{n+1}\sum_{k=0}^{n} \delta_{X_k, T^k\mathbf{e}}$$

satisfies for every $x \in \mathcal{X}$

$$\limsup_{n\to\infty} \frac{1}{n} \log \mathbb{P}_{\mathbf{e},x}(L_n \in F) \leq - \inf_{z\in F} I_{\mathbf{e}}(z)$$

for every closed set F of $\mathcal{M}_1(\mathcal{X} \times E)$, and

$$\liminf_{n\to\infty} \frac{1}{n} \log \mathbb{P}_{\mathbf{e},x}(L_n \in O) \geq - \inf_{z\in O} I_{\mathbf{e}}(z)$$

for every open set O of $\mathcal{M}_1(\mathcal{X} \times E)$.

The existence of such a principle is classical for fixed environment $E = \{\mathbf{e}\}$, finite $\mathcal{X}$, under irreducibility assumption. We refer to Sanov's theorem, see e.g. chapter 6.2 in [DZ98]. We note that the principle can be extended to periodic environments, taking care of the irreducibility. Besides, we are using an analogous result for stationary random environment to get forthcoming Corollary 2, which is due to [S94].

The first question that we tackle now is the mean growth rate of the population. The branching property yields the linearity of the operator $\mu \to m(\mu) = \mathbb{E}_{\mathbf{e},\mu}(Z_1(.))$ for some measurable set A.
In the case of fixed environment, P and N do not depend on $\mathbf{e}$, so m is also fixed and the mean growth rate of the process Z is the limit of $\log \parallel m_n \parallel /n$, with $\parallel . \parallel$ an operator norm. If $\mathcal{X}$ is finite, it yields the Perron-Frobenius eigenvalue under strong irreducibility assumption, with a min-max representation due to Collatz-Wielandt. Krein-Rutman theorem gives an extension to infinite dimension space requiring compactness of the operator m and strict positivity, see also Section 3.4.2.
In the random environment case, it corresponds to the Lyapounov exponent and quenched asymptotic results can be obtained in the case $\mathcal{X}$ is finite [FK60]. Then, the process is a branching process in random environment and we refer to [AK71, K74] for extinction criteria and [C89, T88] for its growth rate.

To go beyond these assumptions and get an interpretation of the growth rate in terms of reproduction-dispersion dynamics, we provide here an other characterization.

[1] It means that $\{\mu \in \mathcal{M}_1(\mathcal{X} \times E) : I(\mu) \leq l\}$ is compact for the weak topology

This is a functional large deviation principle relying on Varadhan's lemma (see also multiplicative ergodicity in [MT09]). It allows to decouple the reproduction and dispersion in the dynamic. Thus, it yields an extension of Theorem 5.3 in [BL12] both for varying environment and infinite state pace $\mathcal{X}$. We refer to this latter article for motivations in ecology, more specifically for metapopulations. The next Corollary then puts in light the dispersion strategy followed by typical individuals of the population for large times.

Theorem 1. *Assume that X satisfies a LDP with good rate function $I_{\mathbf{e}}$ in environment $\mathbf{e}$ and $\log m : \mathcal{X} \times E \to (-\infty, \infty)$ is continuous and bounded. Then, for every $x \in \mathcal{X}$,*

$$\lim_{n\to\infty} \frac{1}{n} \log \mathbb{E}_{\mathbf{e},\delta_x}(Z_n(\mathcal{X})) = \sup_{\mu\in\mathcal{M}_1(\mathcal{X}\times E)} \left\{ \int_{\mathcal{X}\times E} \log(m(x,e))\mu(dxde) - I_{\mathbf{e}}(\mu) \right\} := \varrho_{\mathbf{e}}$$

and

$$M_{\mathbf{e}} := \left\{ \mu \in \mathcal{M}_1(\mathcal{X} \times E) : \int \log(m(x,e))\mu(dxde) - I_{\mathbf{e}}(\mu) = \varrho_{\mathbf{e}} \right\}$$

is compact and non empty.

In particular, $\limsup_{n\to\infty} \frac{1}{n} \log Z_n(\mathcal{X}) \leq \varrho_{\mathbf{e}}$ a.s. The limit can hold only on the survival event. It is the case under classical $N \log N$ moment assumption for finite state space $\mathcal{X}$, see e.g. [LPP95] for one type of individual and fixed environment and [AK71] in random environment. But it is a rather delicate problem when the number of types is infinite.

We assume now $\varrho_{\mathbf{e}} > 0$ and introduce the event

$$\mathcal{S} := \left\{ \liminf_{n\to\infty} \frac{1}{n} \log Z_n(\mathcal{X}) \;\geq\; \varrho_{\mathbf{e}} \right\}.$$

Conditionally on $\mathcal{S}$, we let U_n be an individual uniformly chosen at random in generation n. Let us then focus on its trait frequency up to time n and the associated environment :

$$\nu_n(A) := \frac{1}{n+1} \#\{0 \leq i \leq n : (X_i(U_n), T^i\mathbf{e}) \in A\} \qquad (A \in \mathcal{B}_{\mathcal{X}\times E}).$$

where $X_i(u)$ is the trait of the ancestor of u in generation i. We check now that the support of ν_n converges in probability to $M_{\mathbf{e}}$ on the event $\mathcal{S}$.

Corollary 1. *Under the assumptions of Theorem 1, we further suppose that $\varrho_{\mathbf{e}} > 0$ and $\mathcal{S}$ has positive probability. Then, for every $x \in \mathcal{X}$,*

$$\mathbb{P}_{\mathbf{e},\delta_x}(\nu_n \in \mathcal{C}|\mathcal{S}) \overset{n\to\infty}{\longrightarrow} 0$$

for every closed set $\mathcal{C}$ of $\mathcal{M}_1(\mathcal{X} \times E)$ which is disjoint of $M_{\mathbf{e}}$.

This result deals with the pedigree [JN96, NJ84] or ancestral lineage of a typical individual. It ensures that the trait frequency along the lineage of a typical individual converges to one of the argmax of $\varrho_{\mathbf{e}}$. We are going a bit farther in the next section, with a description of this ancestral lineage via size biased random choice, see in particular Lemma 2.

Let us now specify the theorem for some stationary ergodic environment $\mathcal{E} \in E$. Following [S94], let π be a T invariant ergodic probability, i.e. $\pi \circ T^{-1} = \pi$ and if $A \in \mathcal{B}_E$ satisfies $T^{-1}A = A$, then $\pi(A) \in \{0, 1\}$. Moreover let us require :

Assumption A. There exist a positive integer b, a T invariant subset E' of E and a measurable function $M : E \to [1, \infty)$ such that $\log M \in L^1(\pi)$, $\pi(E') = 1$ and for all $x, y \in \mathcal{X}$, $A \in \mathcal{B}_{\mathcal{X}}$ and $\mathbf{e} \in E'$,

$$P^b(x, \mathbf{e}, A) \leq M(\mathbf{e}) P^b(y, \mathbf{e}, A).$$

We denote by $V_b(\mathcal{X} \times E)$ the set of bounded continuous functions that map $\mathcal{X} \times E$ into $[1, \infty)$ to state the result.

Corollary 2. *Under Assumption A, we further suppose that* $\log m(., \mathcal{E})$ *is* π *a.s. bounded and continuous. Then* π *a.s., for every* $x \in \mathcal{X}$*,*

$$\lim_{n\to\infty} \frac{1}{n} \log \mathbb{E}_{\mathcal{E},\delta_x}(Z_n(\mathcal{X})) = \sup_{\mu \in \mathcal{M}_1(\mathcal{X}\times E)} \left\{ \int \log(m(x,e))\mu(dx,de) - I(\mu) \right\},$$

where I *is defined by*

$$I(\mu) := \sup \left\{ \int_{\mathcal{X}\times E} \log\left(\frac{u(x,e)}{\int_{\mathcal{X}} P(x,e,dy)u(y,Te)} \right) \mu(dx,de) : u \in V_b(\mathcal{X} \times E) \right\}.$$

To prove these results, we need the following lemma, where $\mathcal{B}(S)$ is the set of measurable functions on S. With a slight abuse we note $\mathbb{E}_\nu$ when the process is initiated with one single individual whose trait is distributed as ν. We recall that that $X_i(u)$ is the trait of the ancestor of u in generation i.

Lemma 1. *Let* $F \in \mathcal{B}(\mathcal{X}^k)$ *non-negative. Then, for every* $\nu \in \mathcal{M}_1(\mathcal{X})$*,*

$$\mathbb{E}_{\mathbf{e},\nu}\left(\sum_{|u|=n} F(X_0(u), \ldots, X_n(u)) \right) = \mathbb{E}_{\mathbf{e},\nu}\left(F(X_0, \ldots, X_n) \prod_{i=0}^{n-1} m(X_i, T^i\mathbf{e}) \right).$$

Proof. For every $f_0, \ldots, f_n \in \mathcal{B}(\mathcal{X})$ non-negative, by branching property

$$\mathbb{E}_{\mathbf{e},\nu}\left(\sum_{|u|=n} f_0(X_0(u)) \cdots f_n(X_n(u)) \right)$$
$$= \int \nu(dx_0) f_0(x_0) \int m_1(x_0, \mathbf{e}, dx_1) \mathbb{E}_{T\mathbf{e},\delta_{x_1}} \left(\sum_{|u|=n-1} f_1(X_0(u)) \cdots f_n(X_{n-1}(u)) \right),$$

where

$$m_1(x_0, \mathbf{e}, dx_1) = \mathbb{E}_{\mathbf{e},x_0}\left(\#\{|u| = 1 : X(u) \in dx_1\}\right) = m(x_0, \mathbf{e}) P(x_0, \mathbf{e}, dx_1).$$

So by induction

$$\mathbb{E}_{\mathbf{e},\nu}\left(\sum_{|u|=n} f_0(X_0(u)) \cdots f_n(X_n(u)) \right)$$
$$= \int_{\mathcal{X}^n \times A} \nu(dx_0) f_0(x_0) \prod_{i=0}^{n-1} m(x_i, T^i\mathbf{e}) P(x_i, T^i\mathbf{e}, dx_{i+1}) f_{i+1}(x_{i+1}).$$

It completes the proof. $\square$

Proof of Theorem 1. The previous lemma applied to $F = 1$ ensures that

$$\mathbb{E}_{\mathbf{e},\nu}(Z_n(\mathcal{X})) = \mathbb{E}_{\mathbf{e},\nu}\left(\prod_{i=0}^{n-1} m(X_i, T^i\mathbf{e})\right).$$

Thus

$$\mathbb{E}_{\mathbf{e},\nu}(Z_n(\mathcal{X})) = \mathbb{E}_{\mathbf{e},\nu}\left(\exp\left(n\int_{\mathcal{X}\times E}\log(m(x,e))L^{\mathbf{e}}_{n-1}(dx,de)\right)\right).$$

As $\log m$ is bounded and continuous by assumption, so is

$$\mu \in \mathcal{M}_1(\mathcal{X}\times E) \to \phi(\mu) = \int_{\mathcal{X}\times E}\log(m(x,e))\mu(dx,de).$$

Using the LDP principle satisfied by $L^{\mathbf{e}}_n$, we can apply Varadhan's lemma (see [DZ98] Theorem 4.3.1) to the previous function to get the first part of the Theorem.

Let us now consider a sequence μ_n such that

$$\int_{\mathcal{X}\times E}\log(m(x,e))\mu_n(dxde) - I_{\mathbf{e}}(\mu_n) \overset{n\to\infty}{\longrightarrow} \varrho_{\mathbf{e}}.$$

Then $I_e(\mu_n)$ is upper bounded, which ensures that μ_n belongs to a sublevel set. By Definition 1, such a set is compact so can extract a subsequence μ_{n_k} which converges weakly in $\mathcal{M}(\mathcal{X}, E)$. As I_e is lower semicontinuous, the limit μ of this subsequence satisfies

$$\liminf_{k\to\infty} I_{\mathbf{e}}(\mu_{n_k}) \geq I_{\mathbf{e}}(\mu).$$

Recalling that ϕ is continuous, we get

$$\varrho_{\mathbf{e}} = \lim_{n\to\infty}\left\{\int_{\mathcal{X}\times E}\log(m(x,e))\mu_{\phi(n)}(dxde) - I_{\mathbf{e}}(\mu_{\phi(n)})\right\} \leq \int \log(m(x,e))\mu(dxde) - I_{\mathbf{e}}(\mu)$$

and μ is a maximizer. That ensures that $M_{\mathbf{e}}$ is compact and non empty. $\square$

Proof of Corollary 1. We define for any individual u in generation n

$$\nu_n(u)(A) = \frac{1}{n+1}\sum_{0\leq i\leq n}\delta_{X_i(u)}.$$

Using Lemma 1 with $F(x_0,\ldots,x_n) = \mathbb{1}(\frac{1}{n+1}\sum_{0\leq i\leq n}\delta_{x_i}\in\mathcal{C})$, we have

$$\mathbb{E}_{\mathbf{e},\nu}\left(\#\{u : |u| = n, \nu_n(u)\in\mathcal{C}\}\right) = \mathbb{E}_{\mathbf{e},\nu}\left(\exp\left(n\int_{\mathcal{X}\times E}\log(m(x,e))L^{\mathbf{e}}_{n-1}(dx,de)\right)1_{L^{\mathbf{e}}_n\in\mathcal{C}}\right)$$

Applying again Varadhan's to any bounded continuous function $\phi : \mathcal{M}_1(\mathcal{X}\times E)\to\mathbb{R}$ such that, for every $\mu\in\mathcal{C}$,

$$\phi(\mu) \leq \int_{\mathcal{X}\times E}\log(m(x,e))\mu(dx,de) \tag{1}$$

we get,

$$\limsup_{n\to\infty}\frac{1}{n}\log\mathbb{E}_{\mathbf{e},\nu}\left(\#\{u : |u| = n, F_n(u)\in F\}\right) \quad\leq\quad \sup\{\phi(\mu) - I_{\mathbf{e}}(\mu) : \mu\in\mathcal{M}_1(\mathcal{X}\times E)\}.$$

Let us now check that we can find ϕ such that the right hand side is strictly less than $\varrho_{\mathbf{e}}$. We proceed by contradiction and assume that for every ϕ continuous and bounded which satisfies (1), we have $\sup\{\phi(\mu) - I_{\mathbf{e}}(\mu) : \mu \in \mathcal{M}_1(\mathcal{X} \times E)\} = \varrho_{\mathbf{e}}$. Then using that $I_{\mathbf{e}}$ is a good rate function, we obtain that there exists $\mu(\phi)$ such that $\phi(\mu(\phi)) - I_{\mathbf{e}}(\mu(\phi)) = \varrho_{\mathbf{e}}$, using the same arguments as the end of the previous proof. Recalling that $\mathcal{M}_1(\mathcal{X} \times E)$ can be metrizable by a distance d, we define now $\phi_n(\mu) := -nd(\mu, F) + \int_{\mathcal{X}\times E} \log(m(x,e))\mu(dx,de)$. We use again the compactness of sublevel sets of $I_{\mathbf{e}}$ to extract a sequence $\mu(\phi_{n_k})$ of $\mu(\phi_n)$ which converges to μ_0. Then $\mu_0 \in \mathcal{C} \cap M_{\mathbf{e}}$, which yields the contradiction with $\mathcal{C} \cap M_{\mathbf{e}} = \varnothing$.

Thus we can choose ρ' such that

$$\limsup_{n\to\infty} \frac{1}{n} \log \mathbb{E}_{\mathbf{e},\nu} \left(\#\{u : |u| = n, \nu_n(u) \in \mathcal{C}\}\right) < \varrho' < \varrho_{\mathbf{e}}.$$

Adding that

$$\begin{aligned} \mathbb{P}(\nu_n(U_n) \in F|\mathcal{S}) &\leq \mathbb{E}\left(\#\{u : |u| = n, \nu_n(u) \in \mathcal{C}\}/Z_n(\mathcal{X})|\mathcal{S}\right) \\ &\leq e^{-\varrho' n}\mathbb{E}\left(\#\{u : |u| = n, \nu_n(u) \in \mathcal{C}\}\right)/\mathbb{P}(\mathcal{S}) \end{aligned}$$

for n large enough by definition of $\mathcal{S}$ and that the left hand side goes to 0 ends up the proof. □

Proof of Corollary 2. Under Assumption A, Theorem 3.3 [S94] ensures that there exists a function I which satisfies π a.s. the Definition 1 (uniformly with respect to $x \in \mathcal{X}$). The result is then a direct application of Theorem 1. □

We have given above an expression of the mean growth rate and specified the ancestral lineage of surviving individuals. We are now wondering : does the process grows like its mean when it survives ?
How is the population spread for large times ?

3 Law of large numbers

We consider the mean measure under the environment $\mathbf{e}$:

$$m_n(x, \mathbf{e}, A) := \mathbb{E}_{\mathbf{e},\delta_x}(Z_n(A)) = \mathbb{E}_{\mathbf{e},\delta_x}(\#\{u : |u| = n, X(u) \in A\}) \qquad (A \in \mathcal{B}_{\mathcal{X}}).$$

It yields the mean number of descendant in generation n, whose trait belongs to A, of an initial individual with trait x. By now, we assume that for all $x \in \mathcal{X}, \mathbf{e} \in E$ and $n \geq 0$,

$$m_n(x, \mathbf{e}, \mathcal{X}) < \infty.$$

We define a new family of Markov kernel Q_n by

$$Q_n(x, \mathbf{e}, dy) := m_1(x, \mathbf{e}, dy)\frac{m_{n-1}(y, T\mathbf{e}, \mathcal{X})}{m_n(x, \mathbf{e}, \mathcal{X})}.$$

The fact that $Q_n(x, \mathbf{e}, \mathcal{X}) = 1$ for all $n \in \mathbb{N}, x \in \mathcal{X}, \mathbf{e} \in E$ comes directly from the branching property. We introduce the associated semigroup, more precisely the successive composition of Q_j between the generations i and n :

$$Q_{i,n}(x, \mathbf{e}, A) = Q_{n-i}(x, T^i\mathbf{e}, .) * Q_{n-i-1}(., T^{i+1}\mathbf{e}, .) * \cdots * Q_1(., T^{n-1}\mathbf{e}, .)(A),$$

where we recall the notation $Q(x,.) * Q'(.,.)(A) = \int_{\mathcal{X}} Q(x,dy)Q'(y,A)$ and $\nu(f) = \int_{\mathcal{X}} f(y)\nu(dy)$. The next section links the semigroups m_n and $Q_{0,n}$.

3.1 The auxiliary process and the many-to-one formula

The following many-to-one formula links the expectation of the number of individuals whose trait belongs to A to the probability that the Markov chain associated to the kernel Q_n belongs to A.

Lemma 2. *For all $n \in \mathbb{N}, x \in \mathcal{X}$ and $F \in \mathcal{B}(\mathcal{X}^{n+1})$ non-negative, we have*

$$\mathbb{E}_{\mathbf{e},\delta_x}\left(\sum_{|u|=n} F(X_0(u),\dots,X_n(u))\right) = m_n(x,\mathbf{e},\mathcal{X})\mathbb{E}_{\mathbf{e},x}(F(Y_0^{(n)},\dots,Y_n^{(n)})),$$

where $(Y_i^{(n)} : i = 0,\dots,n)$ is a non-homogeneous Markov chain with transition kernels $(Q_{n-i}(.,T^i\mathbf{e},.) : i = 0,\dots,n-1)$. In particular for each $f \in \mathcal{B}(\mathcal{X})$ non-negative,

$$m_n(x,\mathbf{e},f) = m_n(x,\mathbf{e},\mathcal{X})Q_{0,n}(x,\mathbf{e},f).$$

We note that $m_n(x,\mathbf{e},\mathcal{X})$ is the mean number of individuals in generation n considered in the previous section. Here, combining the branching property and the lemma above yields the following expression of the growth rate:

$$\frac{m_{n+1}(x,\mathbf{e},\mathcal{X})}{m_n(x,\mathbf{e},\mathcal{X})} = \int_{\mathcal{X}} m(y,T^n\mathbf{e})Q_{0,n}(x,\mathbf{e},dy).$$

The many-to-one formula yields the first step of a spine decomposition of the size-biased tree. We have proved that the dynamic of the trait along the spine follows the non-homogeneous Markov chain Y. Going further, we could check that the reproduction of the individuals along the spine is a size biased law and that independent process then grow following the original distribution. Such a decomposition has been firstly achieved for Galton-Watson processes in [LPP95]. We refer to [KLPP97] for an extension to multitype Galton-Watson processes, when the number of types is finite and the environment is fixed, using the eigenvector associated to the maximal eigenvalue of the mean operator. Let us defer to section 3.4.2 some details on this framework and mention [GB03] for continuous time and [G99] for related results in varying environment.
The second part of the Lemma is an extension of the many one-to-one formula for binary tree [G07], Galton-Watson trees [DM10] and Galton-Watson trees in stationary random environments [BH13]. In continuous time, many-to-one formula and formula for forks can been found in [BDMT11]. But these later do not let the reproduction depend on the trait. We refer to [C11, HR12, HR13] for other many-to-one formulas and asymptotic results when reproduction law depend on the trait in some particular cases.
Let us finally mention in this vein the induced random walk eliminating the branching for branching random walk in random environment, see e.g. [CP07a].

Proof. By a telescopic argument :

$$\prod_{i=0}^{n-1} Q_{n-i}(x_i,T^i\mathbf{e},dx_{i+1}) = \frac{m_0(x_n,\mathbf{e},\mathcal{X})}{m_n(x_0,\mathbf{e},\mathcal{X})}\prod_{i=0}^{n-1} m_1(x_i,T^i\mathbf{e},dx_{i+1}).$$

Adding that $m_1(x_i,T^i\mathbf{e},dx_{i+1}) = m(x_i,T^i\mathbf{e})P(x_i,T^i\mathbf{e},dx_{i+1})$, the first part of the lemma is a consequence of Lemma 1. We then deduce the second part by applying the identity obtained to $F(x_0,\dots,x_n) = f(x_n)$. $\square$

Our aim is now to get ride of the expectation and obtain the repartition of the population for large times. We want to derive it from the asymptotic distribution of this auxiliary Markov chain with kernel Q_n and prove a law of large number on the proportions of individuals whose trait belongs to A. Usual methods rely on martingales (see [A00] for infinite number of types) but the assumptions required are not easily fulfilled, at least regarding the motivations from biology and ecology we give in this work. Moreover the generalization to varying environment seems more adapted to the technicals described here. Thus, we are here inspired by ideas and techniques developed in [AK98a, AK98b] using the branching property for suitable large times or that in [G07] relying on L^2 computations. That's why, before preoceeding, we consider the variance of the size of the population and for that purpose we use:

$$\mathcal{V}_i := \{(wa, wb) : a \neq b, |wa| = |wb| = |w| + 1 = i\}.$$

Lemma 3. *Let $\mathbf{e} \in E$, $x \in \mathcal{X}$ and $0 \leq k \leq n$. We have*

$$\begin{aligned}
&\mathbb{E}_{\mathbf{e},\delta_x}(\#\{u, v : |u| = |v| = n,\ |u \wedge v| \geq k\}) \\
&\qquad = m_n(x, \mathbf{e}) + \sum_{i=k+1}^{n} \mathbb{E}_{\mathbf{e},\delta_x}\left(\sum_{(u_1,u_2)\in\mathcal{V}_i} m_{n-i}(X(u_1), T^i\mathbf{e}) m_{n-i}(X(u_2), T^i\mathbf{e})\right).
\end{aligned}$$

In particular, defining

$$V_i(\mathbf{e}, u_1, u_2) = \sup_{k\geq 0} \frac{m_k(X(u_1), T^i\mathbf{e}, \mathcal{X}) m_k(X(u_2), T^i\mathbf{e}, \mathcal{X})}{m_{i+k}(x, \mathbf{e}, \mathcal{X})^2}$$

and assuming that for some sequence $\mathbf{e}_n \in E$,

$$\liminf_{n\to\infty} m_n(x, \mathbf{e}_n, \mathcal{X}) > 0; \quad \sup_{n\geq 0} \sum_{i\geq 1} \mathbb{E}_{\mathbf{e}_n,\delta_x}\left(\sum_{(u_1,u_2)\in\mathcal{V}_i} V_i(\mathbf{e}_n, u_1, u_2)\right) < \infty, \tag{2}$$

then $Z_n(\mathcal{X})/m_n(x, \mathbf{e}_n, \mathcal{X})$ is bounded in $L^2_{\mathbf{e}_n,\delta_x}$.

Proof of Lemma 3. We omit the initial state δ_x in the notations and write $m_n(x, \mathbf{e})$ for $m_n(x, \mathbf{e}, \mathcal{X})$. Using the branching property and distinguishing if the common ancestor of two individuals lives before generation n or in generation n, we have

$$\begin{aligned}
&\mathbb{E}_{\mathbf{e}}(\#\{u, v : |u| = |v| = n,\ |u \wedge v| \geq k\}) \\
&\quad = \mathbb{E}_{\mathbf{e}}\left(\sum_{|u|=|v|=n, |u\wedge v|\geq k} 1\right) \\
&\quad = \mathbb{E}_{\mathbf{e}}(Z_n(\mathcal{X})) + \mathbb{E}_{\mathbf{e}}\left(\sum_{k+1\leq i\leq n} \sum_{\substack{|w|=i-1,\ a\neq b \\ |wa|=|wb|=i}} \sum_{\substack{|u|=n: u\geq wa \\ |v|=n: v\geq wb}} 1\right) \\
&\quad = m_n(x, \mathbf{e}) + \sum_{k+1\leq i\leq n} \mathbb{E}_{\mathbf{e}}\left(\sum_{\substack{|w|=i-1,\ a\neq b \\ |wa|=|wb|=i}} m_{n-i}(X(wa), T^i\mathbf{e}) m_{n-i}(X(wb), T^i\mathbf{e})\right).
\end{aligned}$$

This yields the first part of the Lemma. Then, letting $k = 0$ and dividing by $m_n(x, \mathbf{e})^2$ ensures that

$$\frac{\mathbb{E}_{\mathbf{e}}(Z_n(\mathcal{X})^2)}{m_n(x,\mathbf{e})^2} \leq \frac{1}{m_n(x,\mathbf{e})} + \sum_{i\leq n} \mathbb{E}_{\mathbf{e}} \left(\sum_{(u_1,u_2)\in\mathcal{V}_i} V_i(X(u_1), X(u_2)) \right),$$

which ends up the proof. $\square$

3.2 Branching decomposition

In this part, we focus on the particular case when extinction does not occur and actually assume that the population has a positive growth rate. We have then the following strong law of large numbers, on the following event ensuring the geometric growth of the size of the population :

$$\mathcal{T} := \left\{ \forall n \geq 0,\ Z_n(\mathcal{X}) > 0;\ \liminf_{n\to\infty} \frac{Z_{n+1}(\mathcal{X})}{Z_n(\mathcal{X})} > 1 \right\}.$$

Theorem 2. *Let $\mathbf{e} \in E$ and $f \in \mathcal{B}(\mathcal{X})$ bounded. We assume that there exists a measure ν with finite first moment such that for all $x \in \mathcal{X}, k, l \geq 0$,*

$$\mathbb{P}(N(x, T^k\mathbf{e}) \geq l) \leq \nu[l, \infty). \tag{3}$$

Assume also that there exists a sequence of probability measure μ_n such that

$$\sup_{\substack{\lambda\in\mathcal{M}_1(\mathcal{X})\\ n\geq 0}} \left| Q_{n,n+p}(\lambda, T^n\mathbf{e}, f\circ f_{p+n}) - \mu_{n+p}(f) \right| \longrightarrow 0 \tag{4}$$

as $p \to \infty$. Then,

$$\frac{f_n.Z_n(f)}{Z_n(\mathcal{X})} - \mu_n(f) \stackrel{n\to\infty}{\longrightarrow} 0 \qquad \mathbb{P}_{\mathbf{e},\delta_x} \text{ a.s. on the event } \mathcal{T}. \tag{5}$$

This result is inspired by [AK98a, AK98b]. It extends their approach to the non neutral framework (the reproduction law may depend on the trait) and to time varying environment. It yields a strong law of large numbers relying on the uniform ergodicity of the auxiliary Markov chain $Q_{i,n}$. The restriction to the event where the size of the population has positive growth rate will be relaxed in the next part using L^2 assumptions. We here simply mention that the previous convergence can be stated on the (a priori larger) event

$$\mathcal{T}' := \left\{ \forall n \geq 0,\ Z_n(\mathcal{X}) > 0;\ \liminf_{n\to\infty} \frac{1}{n}\log(Z_n(\mathcal{X})) > 0 \right\}.$$

This extension may be useful for applications in random environmnent, when subcritical environments occur i.o. To prove this, one need to replace (3) by a stronger assumption (uniform L^2 bound) and adapt the Lemma below.

We first state a law of large numbers, which is being used several time. It is an easy extension of Lemma 1 in [AK98a], which itself is proved using [K72].

Lemma 4. *Let $\{\mathcal{F}_n\}_0^\infty$ be a filtration contained in $(\Omega, \mathcal{B}, \mathbb{P})$. Let $\{X_{n,i} : n, i \geq 1\}$ be r.v. such that for each n, conditionally on $\mathcal{F}_n$, $\{X_{n,i} : i \geq 1\}$ are centered independent r.v. Let $\{N_n : n \geq 1\}$ be non-negative integer valued r.v. such that for each n, N_n is $\mathcal{F}_n$ measurable.*
We assume that there exists a random measure ν with finite first moment such that

$$\forall t > 0, \qquad \sup_{n,i \geq 1} \mathbb{P}(|X_{n,i}| > t | \mathcal{F}_n) \leq \nu(t, \infty) \quad a.s.$$

Then

$$\frac{1}{N_n} \sum_{i=1}^{N_n} X_{n,i} \overset{n\to\infty}{\longrightarrow} 0$$

a.s. on the event $\{\forall n \geq 0 : N_n > 0, \ \liminf_{n\to\infty} N_{n+1}/N_n > 1\}$.

Proof. The proof can be simply adapted from the proof of Lemma 1 in [AK98a]. For any $\delta > 0$, $n_0 \geq 1$ and $l > 1$, we define

$$A_n := \left\{ \left| \frac{1}{N_n} \sum_{i=1}^{N_n} X_{n,i} \right| > \delta; \ \forall k = n_0, \ldots, n : \frac{N_k}{N_{k-1}} \geq l \right\}$$

and prove similarly that $\sum_{n \geq n_0} \mathbb{P}(A_n | \mathcal{F}_n) < \infty$. This yields the expected a.s. convergence on the event $\{\forall n \geq n_0, \ N_{n+1}/N_n \geq l\}$ by conditional Borel Cantelli Lemma [Chow and Teicher, 1988, p249] and the result follows by monotone limit. ☐

We use this lemma to prove the following result.

Lemma 5. *Under the assumptions of Theorem 2, we have*

$$\frac{1}{Z_{n+p}(\mathcal{X})} \sum_{|u|=n} m_p(X(u), T^n\mathbf{e}, \mathcal{X}) \overset{n\to\infty}{\longrightarrow} 1 \quad \mathbb{P}_{\mathbf{e},\delta_x} \ a.s \ on \ \mathcal{T}.$$

Proof. The branching property gives a natural decomposition of the population in generation $n + p$:

$$Z_{n+p}(\mathcal{X}) = \sum_{|u|=n} Z_p^{(u)}(\mathcal{X}),$$

where $Z^{(u)}$ is the branching Markov chain whose root is the individual u and whose environment is $T^n\mathbf{e}$. Moreover,

$$Z_{n+p}(\mathcal{X}) - \sum_{|u|=n} m_p(X(u), T^n\mathbf{e}, \mathcal{X}) = \sum_{|u|=n} \left[Z_p^{(u)}(\mathcal{X}) - m_p(X(u), T^n\mathbf{e}, \mathcal{X}) \right] = Z_n(\mathcal{X}) \epsilon_{n,p},$$

where

$$\epsilon_{n,p} := \frac{1}{Z_n(\mathcal{X})} \sum_{|u|=n} X_{p,u}^{(n)}, \qquad X_{p,u}^{(n)} := Z_p^{(u)}(\mathcal{X}) - m_p(X(u), T^n\mathbf{e}, \mathcal{X}).$$

We note that $(X_{p,u}^{(n)} : |u| = n)$ are independent conditionally on $\mathcal{F}_n = \sigma(X(v) : |v| \leq n)$, $\mathbb{E}(X_{p,u}^{(n)}) = 0$ and $|X_{p,u}^{(n)}| \leq |Z_p^{(u)}(\mathcal{X})| + m_p(X(u), T^n\mathbf{e}, \mathcal{X})$, so that the stochastic domination assumption (3) ensures that there exists a measure with finite first moment ν' such that

$$\sup_{u \in \mathbb{T}} \mathbb{P}_{\mathbf{e},\delta_x}(|X_{p,u}^{(n)}| > t | \mathcal{F}_{|u|}) \leq \nu'(t, \infty),$$

where we recall that $\mathbb{T}$ is the set of all individuals. We can then apply the law of large number of Lemma 5 to get that for every $p \geq 0$, $\epsilon_{n,p} \to 0$ a.s. on the event $A_{n_0,l}$, as $n \to \infty$. Recalling that $Z_{n+p}(\mathcal{X}) \geq Z_n(\mathcal{X})$ for n large enough ends up the proof. $\square$

We can now prove the Theorem.

Proof of Theorem 2. We use again the branching decomposition in generation n to write

$$\begin{aligned}
&\left| \frac{f_{n+p}.Z_{n+p}(f)}{Z_{n+p}(\mathcal{X})} - \mu_{n+p}(f) \right| \\
&\quad = \left| \frac{1}{Z_n(\mathcal{X})} \sum_{|u|=n} \frac{Z_n(\mathcal{X})}{Z_{n+p}(\mathcal{X})} f_{n+p}.Z_p^{(u)}(f) - \mu_{n+p}(f) \right| \\
&\quad \leq \left| \sum_{|u|=n} \frac{X_{u,n,p}}{Z_n(\mathcal{X})} \right| + \left| \sum_{|u|=n} \frac{Y_{u,n,p}}{Z_{n+p}(\mathcal{X})} \right| + \mu_{n+p}(f) \left| \sum_{|u|=n} \frac{m_p(X(u), T^n\mathbf{e}, \mathcal{X})}{Z_{n+p}(\mathcal{X})} - 1 \right|,
\end{aligned}$$

where

$$X_{u,n,p} = \frac{Z_n(\mathcal{X})}{f_{n+p}.Z_{n+p}(\mathcal{X})} \left[f_{n+p}.Z_p^{(u)}(f) - m_p(X(u), T^n\mathbf{e}, f \circ f_{n+p}) \right]$$

and

$$Y_{u,n,p} = m_p(X(u), T^n\mathbf{e}, \mathcal{X})[Q_{0,p}(X(u), T^n\mathbf{e}, f \circ f_{n+p}) - \mu_{n+p}(f)].$$

We want to prove that these quantities go to zero. First we note that

$$\begin{aligned}
X_{u,n,p} &\leq f_{n+p}.Z_p^{(u)}(f) + m_p(X(u), T^n\mathbf{e}, f \circ f_{n+p}) \\
&\leq \| f \|_\infty \, [Z_p^{(u)}(\mathcal{X}) + m_p(X(u), T^n\mathbf{e}, \mathcal{X})],
\end{aligned}$$

so that Assumption (3) ensures that the r.v. $X_{u,n,p}$ are stochastically dominated. Then, we can apply the law of large numbers of Lemma 5 and get that for each $p \geq 0$, $\sum_{|u|=n} X_{u,n,p}/Z_n(\mathcal{X})$ tends to zero as $n \to \infty$.
Moreover the many-to-one formula (Lemma 2) ensures that

$$Y_{u,n,p} = m_p(X(u), T^n\mathbf{e}, \mathcal{X})[Q_{0,p}(X(u), T^n\mathbf{e}, f \circ f_{n+p}) - \mu_{n+p}(f)] \leq m_p(X(u), T^n\mathbf{e}, \mathcal{X})M_p$$

where

$$M_p := \sup_{n \geq 0} |Q_{0,p}(X(u), T^n\mathbf{e}, f \circ f_{n+p}) - \mu_{n+p}(f)|.$$

Combining these results, we get

$$\begin{aligned}
&\limsup_{n\to\infty} \left| \frac{f_{n+p}.Z_{n+p}(f)}{Z_{n+p}(\mathcal{X})} - \mu_{n+p}(f) \right| \\
&\quad \leq M_p \limsup_{n\to\infty} \sum_{|u|=n} \frac{m_p(X(u), T^n\mathbf{e}, \mathcal{X})}{Z_{n+p}(\mathcal{X})} \\
&\qquad + \| f \|_\infty \limsup_{n\to\infty} \left| \sum_{|u|=n} \frac{m_p(X(u), T^n\mathbf{e}, \mathcal{X})}{Z_{n+p}(\mathcal{X})} - 1 \right| \leq M_p.
\end{aligned}$$

by means of Lemma 5. Using now that $M_p \to 0$ as $p \to \infty$ by (4), we have

$$\limsup_{p\to\infty} \limsup_{n\to\infty} \left| \frac{Z_{n+p}(f)}{f_{n+p}.Z_{n+p}(\mathcal{X})} - \mu_{n+p}(f) \right| = 0$$

and the proof is complete. $\square$

3.3 L^2 convergence

In this section, we state weak and strong law of large numbers by combining L^2 computations, the ergodicity of the auxiliary Markov chain Y and the position of the most recent common ancestor of the individuals in generation n.
We recall the notations $Q(\lambda, \mathbf{e}, f)(x) = \int_{\mathcal{X}^2} \lambda(dx)Q(x, \mathbf{e}, dy)f(y)$ and $\mathcal{B}(\mathcal{X})$ for the set of measurable functions from $\mathcal{X}$ to $\mathbb{R}$. We note $\mathcal{B}_b(\mathcal{X})$ the set of measurable functions from $\mathcal{X}$ to $\mathbb{R}$, which are bounded by a same constant $b \geq 0$.

The main assumption we are using concern the ergodic behavior of the time non-homogeneous auxiliary Markov chain Y.

Assumption 1. Let $\mathbf{e}_n \in E$, $\mathcal{F} \subset \mathcal{B}(\mathcal{X})$, $f_n \in \mathcal{B}(\mathcal{X})$ and $\mu_n \in \mathcal{M}_1(\mathcal{X})$ for each $n \in \mathbb{N}$.
(a) For all $\lambda \in \mathcal{M}_1(\mathcal{X})$ and $i \in \mathbb{N}$,

$$\sup_{f\in\mathcal{F}} \left|Q_{i,n}(\lambda, \mathbf{e}_n, f \circ f_n) - \mu_n(f)\right| \stackrel{n\to\infty}{\longrightarrow} 0.$$

(b) For every $k_n \leq n$ such that $n - k_n \to \infty$,

$$\sup_{\lambda\in\mathcal{M}_1(\mathcal{X}), f\in\mathcal{F}} \left|Q_{k_n,n}(\lambda, \mathbf{e}_n, f \circ f_n) - \mu_n(f)\right| \stackrel{n\to\infty}{\longrightarrow} 0.$$

We note that we can set $\mu_n(f) := Q_{0,n}(\lambda, \mathbf{e}_n, f \circ f_n)$ in this assumption, for any fixed $\lambda \in \mathcal{M}_1(\mathcal{X})$. The second assumption (uniform ergodicity) clearly implies the first one. Sufficient conditions will be given in the applications. In particular, they are linked to Harris ergodic theorem and more specifically they will be formulated in terms of Doeblin and Lyapounov type conditions. The function f_n is bound to make the process ergodic if it is not originally. We have for example in mind the case when the auxiliary Markov chain X_n satisfies a central limit theorem, $f_n(x) = (x - a_n)/b_n$ and $f(X_n)$ converges to the same distribution whatever the initial value X_0 is. Such convergence hold for example for branching random walks.

We consider now the genealogy of the population and the time of the most recent common ancestor of two individuals chosen uniformly.

Assumption 2. (a) For every $\epsilon > 0$, there exists $K \in \mathbb{N}$, such that for n large enough,

$$\frac{\mathbb{E}_{\mathbf{e}_n,\delta_x}(\#\{u, v : |u| = |v| = n,\ |u \wedge v| \geq K\})}{m_n(x, \mathbf{e}_n, \mathcal{X})^2} \leq \epsilon. \tag{6}$$

Moreover there exists $C_i \in \mathcal{B}(\mathcal{X}^2)$ such that for all $i \in \mathbb{N}, x, y \in \mathcal{X}$,

$$\sup_{n\geq i} \frac{m_{n-i}(y, T^i\mathbf{e}_n, \mathcal{X})}{m_n(x, \mathbf{e}_n, \mathcal{X})} \leq C_i(x, y), \quad \text{with } \mathbb{E}\left(\max\{C_i(x, X(w))^2 : |w| = i + 1\}\right) < \infty.$$

(b) For every $K \in \mathbb{N}$,

$$\frac{\mathbb{E}_{\mathbf{e}_n,\delta_x}(\#\{u, v : |u| = |v| = n,\ |u \wedge v| \geq n - K\})}{m_n(x, \mathbf{e}_n, \mathcal{X})^2} \stackrel{n\to\infty}{\longrightarrow} 0. \tag{7}$$

Moreover,

$$\sup_{n\in\mathbb{N}} \mathbb{E}_{\mathbf{e}_n,\delta_x}(Z_n(\mathcal{X})^2)/m_n(x, \mathbf{e}_n, \mathcal{X})^2 < \infty.$$

These expressions can be rewritten in terms of normalized variance of $Z_n(\mathcal{X})$ and more tractable sufficient assumptions can be specified, using Lemma 3, see also the applications below. We also observe that these assumptions require that each reproduction law involved in the dynamic has a finite second moment. Moreover $m_n(x, \mathbf{e}_n, \mathcal{X})$ has to go to $+\infty$.

The assumption (6) says that the common ancestor is at the beginning of the genealogy. It is the case for Galton-Watson trees, branching processes in random environment and many others "regular trees". The assumption (7) says that the common ancestor is not at the end of the genealogy, so it is weaker. For a simple example where (6) is fulfilled but (7) is not, one can consider the tree T_n which is composed by a single individual until the generation $n - k_n$ and equal to the binary tree between the generations $n - k_n$ and n, with $k_n \to \infty$. One can also construct examples of branching Markov chain with time homogeneous reproduction. As an hint, we mention the degenerated case when the tree is formed by a spine where each individual has exactly one offsprings, but the individuals of the spine which have two offsprings. More generally, such a genealogy may arise by considering increasing Markov chains and increasing mean reproduction $m(x, \mathbf{e})$ with respect to $x \in \mathcal{X}$.

Theorem 3 (Weak LLN)**.** *Let $\mathbf{e}_n \in E$, $x \in \mathcal{X}$, $f_n : \mathcal{X} \to \mathcal{X}$ and F be a subset of $\mathcal{B}(\mathcal{X})$ such that $\sup_{f\in\mathcal{F}} \parallel f \parallel_\infty < \infty$.*

We assume either that Assumptions 1(a) and 2(a) hold or that Assumptions 1(b) and 2(b) hold. Then, uniformly for $f \in \mathcal{F}$,

$$\frac{f_n.Z_n(f) - \mu_n(f)Z_n(\mathcal{X})}{m_n(x, \mathbf{e}_n, \mathcal{X})} \overset{n\to\infty}{\longrightarrow} 0 \tag{8}$$

in $L^2_{\mathbf{e}_n,\delta_x}$ and for all $\epsilon, \eta > 0$,

$$\mathbb{P}_{\mathbf{e}_n,\delta_x}\left(\left|\frac{f_n.Z_n(f)}{Z_n(\mathcal{X})} - \mu_n(f)\right| \geq \eta\ ;\ Z_n(\mathcal{X})/m_n(x, \mathbf{e}_n, \mathcal{X}) \geq \epsilon\right) \overset{n\to\infty}{\longrightarrow} 0.$$

We first note that $f_n.Z_n(\mathbb{1}(A))/Z_n(\mathcal{X})$ is the proportion of individuals in generation n whose trait belongs to $f_n^{-1}(A)$. The assumptions require either weak ergodicity and early separation of lineages or strong ergodicity and non-late separation of lineages.

We refer to the next section for various applications.

We also mention that the Theorem holds if $f_n : \mathcal{X} \to \mathcal{X}'$ and can be extended to unbounded f with domination assumptions following [G07]. Finally, let us mention that the a.s. convergence may fail in the theorem above, even in the field of applications we can have in mind. One can think for example of an underlying genealogical tree growing very slowly and each individual is attached with i.i.d. random variable, as appears in tree indexed random walks, using the rate function of the associated random walk.

Proof. Let us prove the first part of the Theorem under Assumptions 1(a) and 2(a). In the whole proof, x is fixed and we omit δ_x in the notation of the probability and of the expectation. For convenience, we also write $m(x, \mathbf{e}_n) := m(x, \mathbf{e}_n, \mathcal{X})$, $b := \sup_{f\in\mathcal{F}} \parallel f \parallel_\infty$ and denote

$$g_n(x) := f(f_n(x)) - \mu_n(f).$$

We compute for $K \geq 1$,

$$\begin{aligned}
&\mathbb{E}_{\mathbf{e}_n}\left(Z_n(g_n)^2\right)\\
&= \mathbb{E}_{\mathbf{e}_n}\left(\sum_{|u|=|v|=n} g_n(X(u))g_n(X(v)\right)\\
&= \mathbb{E}_{\mathbf{e}_n}\left(\sum_{\substack{|u|=|v|=n\\|u\wedge v|<K}} g_n(X(u))g_n(X(v))\right) + \mathbb{E}_{\mathbf{e}_n}\left(\sum_{\substack{|u|=|v|=n\\|u\wedge v|\geq K}} g_n(X(u))g_n(X(v))\right)
\end{aligned}$$

The second term of the right hand side is smaller than

$$2 \parallel f \parallel_\infty^2 \mathbb{E}(\#\{|u| = |v| = n : |u \wedge v| > K\}) \leq 2b^2 m(x, \mathbf{e}_n)^2 . \epsilon_{K,n},$$

where $\limsup_{n\to\infty} \epsilon_{K,n} \to 0$ as $K \to \infty$ using the first part of Assumption 2(a). So we just deal with the first term and consider $i = 1, \ldots, K$. Thanks to the branching property,

$$\begin{aligned}
\mathbb{E}_{\mathbf{e}_n}\left(\sum_{\substack{|u|=|v|=n\\|u\wedge v|=i-1}} g_n(X(u))g_n(X(v))\right) &= \mathbb{E}_{\mathbf{e}_n}\left(\sum_{\substack{|w|=i-1\\|wa|=|wb|=i}} \sum_{\substack{|u|=n\\u\geq wa}} \sum_{\substack{|v|=n\\v\geq wb}} g_n(X(u))g_n(X(v))\right)\\
&= \mathbb{E}_{\mathbf{e}_n}\left(\sum_{\substack{|w|=i-1\\|wa|=|wb|=i}} R_{i,n}(X(wa))R_{i,n}(X(wb))\right),
\end{aligned}$$

where the many-to-one formula of Lemma 2 allows us to write

$$R_{i,n}(x) := \mathbb{E}_{T^i\mathbf{e}_n,\delta_x}\left(\sum_{|u|=n-i} g_n(X(u))\right) = m_{n-i}(x, T^i\mathbf{e}_n)Q_{0,n-i}(x, T^i\mathbf{e}_n, g_n). \qquad (9)$$

Then

$$\begin{aligned}
&m(x, \mathbf{e}_n, \mathcal{X})^{-2}\mathbb{E}_{\mathbf{e}_n}\left(\sum_{\substack{|u|=|v|=n\\|u\wedge v|\leq K}} g_n(X(u))g_n(X(v))\right)\\
&= \mathbb{E}_{\mathbf{e}_n}\left(\sum_{\substack{i\leq K,|w|=i-1\\|wa|=|wb|=i}} F_{i,n}(wa)F_{i,n}(wb)\frac{m_{n-i}(X(wa), T^i\mathbf{e}_n)m_{n-i}(X(wb), T^i\mathbf{e}_n, \mathcal{X})}{m_n(x, \mathbf{e}_n)^2}\right).
\end{aligned}$$

and Assumption 1(a) ensures that

$$F_{i,n}(u) := R_{i,n}(X(u))/m_{n-i}(X(u), T^i\mathbf{e}_n)$$

goes to 0 a.s. for each $i \in \mathbb{N}$ and $u \in \mathbb{T}$ such that $|u| = i$. Moreover this convergence is uniform for $f \in \mathcal{F}$. Adding that $F_{i,n}$ is bounded by b, we have

$$\begin{aligned}
&F_{i,n}(wa)F_{i,n}(wb)\frac{m_{n-i}(X(wa), T^i\mathbf{e}_n)m_{n-i}(X(wb), T^i\mathbf{e}_n)}{m_n(x, \mathbf{e}_n)^2}\\
&\qquad\qquad \leq b^2 \sup_n \frac{m_{n-i}(X(wa), T^i\mathbf{e}_n)}{m_n(x, \mathbf{e}_n)} . \sup_n \frac{m_{n-i}(X(wb), T^i\mathbf{e}_n)}{m_n(x, \mathbf{e}_n)}.
\end{aligned}$$

By bounded convergence, the second part of Assumption 2(a) ensures that $Z_n(g_n)/m_n(x, \mathbf{e}_n)$ converges to 0 in $L^2_{\mathbf{e}_n}$ uniformly for $f \in \mathcal{F}$. It ends up the proof of (8) under Assumptions 1(a) and 2(a) .

The proof of (8) under Assumptions 1(b) and 2(b) is almost the same, replacing K by $n - k_n$ with $k_n \to \infty$. Indeed, Assumption 2(b) ensures that there exists $k_n \to \infty$ such that

$$\frac{\mathbb{E}_{\mathbf{e}_n,\delta_x}\left(\#\{|u| = |v| = n : |u \wedge v| > n - k_n\}\right)}{m_n(x, \mathbf{e}_n)^2} \overset{n\to\infty}{\longrightarrow} 0,$$

whereas

$$\mathbb{E}_{\mathbf{e}_n}\left(\sum_{\substack{i \le n-k_n, |w|=i-1 \\ |wa|=|wb|=i}} F_{i,n}(wa)F_{i,n}(wb)\frac{m_{n-i}(X(wa), T^i\mathbf{e}_n)m_{n-i}(X(wb), T^i\mathbf{e}_n)}{m_n(x, \mathbf{e}_n)^2}\right)$$

$$\le \left(\sup_{n-i \ge k_n, x\in\mathcal{X}} F_{i,n}(x)\right)^2 \frac{\mathbb{E}(Z_n(\mathcal{X})^2)}{m_n(x, \mathbf{e}_n)^2}.$$

Assumption 1(b) ensures that $\sup_{n-i\ge k_n, x\in\mathcal{X}} F_{i,n}(x) \to 0$ as $k_n \to \infty$ and the second part of Assumption 2(b) ensures that $\mathbb{E}_{\mathbf{e}_n}(Z_n(\mathcal{X})^2)/m_n(x, \mathbf{e}_n)^2$ is bounded. The conclusion is thus the same.

The proof of the last part of the Theorem comes simply from Cauchy-Schwarz inequality :

$$\mathbb{E}_{\mathbf{e}_n}\left(\mathbb{1}_{Z_n(\mathcal{X})/m(x,\mathbf{e}_n)\ge\epsilon}\left[\frac{f_n.Z_n(f)}{Z_n(\mathcal{X})} - \mu_n(f)\right]\right)^2$$
$$\le \mathbb{E}_{\mathbf{e}_n}\left(\frac{m_n(x, \mathbf{e}_n)^2}{Z_n(\mathcal{X})^2} 1_{Z_n(\mathcal{X})/m_n(x,\mathbf{e}_n)\ge\epsilon}\right)\mathbb{E}_{\mathbf{e}_n}\left(\left[\frac{f_n.Z_n(f) - Z_n(\mathcal{X})\mu_n(f)}{m_n(x, \mathbf{e}_n)}\right]^2\right).$$

The first term of the right-hand side is bounded with respect to n. So applying the first part of the Theorem to the second term and using Markov inequality ends up the proof. □

We give now a strong law of large numbers.

Theorem 4 (Strong LLN)**.** *Let $\mathbf{e} \in E$, $x \in \mathcal{X}$ and $f \in \mathcal{B}_b(\mathcal{X})$.*
Assume that

$$\liminf_{n\to\infty} m_n(x, \mathbf{e}, \mathcal{X}) > 0; \quad \sum_{i\ge1} \mathbb{E}_{\mathbf{e},\delta_x}\left(\sum_{(u_1,u_2)\in\mathcal{V}_i} V_i(\mathbf{e}, u_1, u_2)\right) < \infty, \tag{10}$$

and that there exists a sequence of probability measure μ_n on $\mathcal{X}$ such that

$$\sup_{i\in\mathbb{N}} \sum_{n\ge i} \sup_{\lambda\in\mathcal{M}_1(\mathcal{X})} \left|Q_{i,n}(\lambda, T^i\mathbf{e}, f \circ f_n) - \mu_n(f)\right|^2 < \infty. \tag{11}$$

Then $Z_n(\mathcal{X})/m_n(x, \mathbf{e}, \mathcal{X})$ is bounded in $L^2_{\mathbf{e},\delta_x}$ and

$$\frac{f_n.Z_n(f) - \mu_n(f)Z_n(\mathcal{X})}{m_n(x, \mathbf{e}, \mathcal{X})} \overset{n\to\infty}{\longrightarrow} 0 \qquad \mathbb{P}_{\mathbf{e},\delta_x} \ a.s.$$

The first assumption is related to the genealogy of the population and the second one is linked to the ergodic property of the auxiliary Markov chain Y. Both assumptions are stronger that their counterpart of the previous theorem.
As $Z_n(\mathcal{X})/m_n(x,\mathbf{e},\mathcal{X})$ is bounded in $L^2_{\mathbf{e}}$, the event $\{Z_n(\mathcal{X})/m_n(x,\mathbf{e},\mathcal{X}) \geq \epsilon\}$ is positive for ϵ small enough and every $n \geq 1$. On this event, we note that $f_n.Z_n(f)/Z_n(\mathcal{X}) - \mu_n(f) \to 0$ goes to 0 as $n \to \infty$, as in Theorem 2.

Proof. The fact that $Z_n(\mathcal{X})/m_n(x,\mathbf{e},\mathcal{X})$ is bounded in $L^2_{\mathbf{e},\delta_x}$ comes directly from the second part of Lemma 3. To get the a.s. convergence, we prove that

$$\mathbb{E}_{\mathbf{e}}\left(\sum_{n\geq 1}\left[\frac{f_n.Z_n(f)-\mu_n(f)Z_n(\mathcal{X})}{m_n(x,\mathbf{e})}\right]^2\right) < \infty.$$

For that purpose, we use the notations of the proof of the previous Theorem, in particular

$$g_n(x) := f(f_n(x)) - \mu_n(f)$$

and we are inspired by L^2 computations for Markov chain indexed by trees, see e.g. [G07]. Using Fubini inversion, the branching property and (9), we have

$$\begin{aligned}
&\sum_{n\geq 0} m_n(x,\mathbf{e})^{-2}\mathbb{E}_{\mathbf{e}}(Z_n(g_n)^2)\\
&\quad= \mathbb{E}_{\mathbf{e}}\left(\sum_{n\in\mathbb{N}}\sum_{|u|=|v|=n} m_n(x,\mathbf{e})^{-2} g_n(X(u))g_n(X(v))\right)\\
&\quad= \mathbb{E}_{\mathbf{e}}\left(\sum_{n\in\mathbb{N}}\sum_{i\leq n}\sum_{\substack{|u|=|v|=n\\ |u\wedge v|=i}} m_n(x,\mathbf{e})^{-2} g_n(X(u))g_n(X(v))\right)\\
&\quad= \mathbb{E}_{\mathbf{e}}\left(\sum_{\substack{n\in\mathbb{N}\\ i\leq n}}\sum_{\substack{|w|=i-1\\ |wa|=|wb|=i}}\sum_{\substack{|u|=n:u\geq wa\\ |v|=n:v\geq wb}} m_n(x,\mathbf{e})^{-2} g_n(X(u))g_n(X(v))\right)\\
&\qquad\quad+\mathbb{E}_{\mathbf{e}}\left(\sum_{n\in\mathbb{N},|u|=n} m_n(x,\mathbf{e})^{-2} g_n(X(u))^2\right).
\end{aligned}$$

Then

$$
\begin{aligned}
&\sum_{n\geq 0} m_n(x,\mathbf{e})^{-2}\mathbb{E}_{\mathbf{e}}(Z_n(g_n)^2) \\
&\leq \mathbb{E}_{\mathbf{e}}\left(\sum_{\substack{n\in\mathbb{N}\\ i\leq n}}\sum_{\substack{|w|=i-1\\ |wa|=|wb|=i}} \frac{m_{n-i}(X(wa),T^i\mathbf{e})m_{n-i}(X(wb),T^i\mathbf{e})}{m_n(x,\mathbf{e})^2}F_{i,n}(X(wa))F_{i,n}(X(wb))\right) \\
&\qquad +2 \parallel g_n \parallel_\infty \mathbb{E}_{\mathbf{e}}\left(\sum_{n\in\mathbb{N}} m_n(x,\mathbf{e})^{-2}Z_n(\mathcal{X})\right) \\
&\leq \mathbb{E}_{\mathbf{e}}\left(\sum_{\substack{i\in\mathbb{N},|w|=i-1\\ |wa|=|wb|=i}} V_i(X(wa),X(wb))H_i\right) + b\sum_{n\in\mathbb{N}} m_n(x,\mathbf{e})^{-1},
\end{aligned}
$$

where $b := (2 \parallel f \parallel_\infty)^2$ and

$$
H_i = \sup_{y,z}\sum_{n\geq i} F_{i,n}(y)F_{i,n}(z), \qquad V_i(x_0,x_1) = \sup_{n\geq i}\frac{m_{n-i}(x_0,T^i\mathbf{e})m_{n-i}(x_1,T^i\mathbf{e})}{m_n(x,\mathbf{e})^2}.
$$

The assumptions (10) and (11) ensure that

$$
\begin{aligned}
&\sum_{n\geq 0} m_n(x,\mathbf{e},\mathcal{X})^{-2}\mathbb{E}_{\mathbf{e}}(Z_n(g_n)^2) \leq \\
&\qquad b\sum_{n\geq 0} m_n(x,\mathbf{e})^{-1} + \sup_{i\in\mathbb{N}} H_i.\sum_{i\in\mathbb{N}}\mathbb{E}\left(\sum_{(u_1,u_2)\mathcal{V}_i} V_i(X(u_1),X(u_2))\right) < \infty.
\end{aligned}
$$

Then, $Z_n(g_n)/m_n(x,\mathbf{e}) \to 0$ a.s., which completes the proof. □

3.4 Applications based on the spectral theory of the mean operator and geometric ergodicity

We aim now at providing sufficient tractable conditions to apply the law of large numbers stated in the previous sections. We focus here on strong law of large numbers we can derive from Theorem 4 using the ergodicity of the (one-dimensional) auxiliary Markov chain. Analogous results could be stated using Theorem 2 but seem to be more restrictive when considering applications in varying environment. More precisely, we first focus on the neutral case and the separation of the growth term and the dynamic of the trait is very natural. We then use classical spectral theory and the spectral gap of mean operator to state a result which can be applied in the time homogeneous framework. Finally we use focus on the geometric Harris ergodicity of the auxiliary Markov chain and check that the assumptions of the previous Theorems can be fulfilled and applied to non-neutral branching processes in varying environment.

We refer to the next subsection for additional applications. Let us begin with a lemma, which provides a convenient sufficient condition for Assumption 2. It will be be satisfied in the following applications. We recall that it ensures that the common ancestor of two individuals chosen independently lives at the beginning of the tree.

Lemma 6. *Assume that there exists* $(C(\mathbf{e}) : \mathbf{e} \in E)$ *such that for all* $x, y \in \mathcal{X}$ *and* $n \geq 0$*,*

$$m_n(y, \mathbf{e}, \mathcal{X}) \leq C(\mathbf{e}) m_n(x, \mathbf{e}, \mathcal{X}) \tag{12}$$

and a sequence $\mathbf{e}_n \in E$ *such that*

$$\sup_{n \geq 0} \sum_{i=1}^{n} \frac{1 \wedge D(x, T^{i-1}\mathbf{e}_n)}{m_i(x, \mathbf{e}_n, \mathcal{X})} < \infty, \tag{13}$$

where

$$\sigma(\mathbf{e}) := \sup_{y \in \mathcal{X}} \mathbb{E}(N(y, \mathbf{e})^2), \qquad D(x, \mathbf{e}) := \frac{\sigma(\mathbf{e}) C(\mathbf{e})}{m(x, \mathbf{e})} C(T\mathbf{e})^2.$$

Then (2) holds and $Z_n(\mathcal{X})/m_n(x, \mathbf{e}_n, \mathcal{X})$ *is bounded in* $L^2_{\mathbf{e}_n, \delta_x}$*. Moreover, if* $\mathbf{e}_n = \mathbf{e}$*, then (10) also holds.*

Proof. Using the branching property in generation i and (12), we have for all $y \in \mathcal{X}$,

$$m_{i+k}(x, \mathbf{e}) \geq m_i(x, \mathbf{e}) C(T^i\mathbf{e})^{-1} m_k(y, T^i\mathbf{e}). \tag{14}$$

Then

$$V_i(\mathbf{e}, u_1, u_2) \leq \frac{C(T^i\mathbf{e})^2}{m_i(x, \mathbf{e}, \mathcal{X})^2}$$

and

$$\mathbb{E}_{\mathbf{e}_n}\left(\sum_{(u_1, u_2) \in \mathcal{V}_i} V_i(u_1, u_2) \right) \leq \mathbb{E}_{\mathbf{e}_n}(\mathcal{V}_i) \frac{C(T^i\mathbf{e})^2}{m_i(x, \mathbf{e}, \mathcal{X})^2}.$$

Adding that

$$\mathbb{E}_{\mathbf{e}}(\mathcal{V}_i) \leq m_{i-1}(x, \mathbf{e}, \mathcal{X}) \sigma(T^{i-1}\mathbf{e})$$

and using again (14) with $k = 1$, we get

$$\mathbb{E}_{\mathbf{e}_n, \delta_x}\left(\sum_{(u_1, u_2) \in \mathcal{V}_i} V_i(u_1, u_2) \right) \leq \frac{\sigma(T^{i-1}\mathbf{e}) C(T^{i-1}\mathbf{e}_n)}{m(x, T^{i-1}\mathbf{e})} \frac{C(T^i\mathbf{e}_n)^2}{m_i(x, \mathbf{e}_n, \mathcal{X})}.$$

Thus, (2) and (10) hold. Applying Lemma 3 ends up the proof. □

3.4.1 Weak law of large numbers along branching trees (neutral case)

First, we consider the neutral case, which means that the reproduction law of an individual does not depend on their trait. Thus the underlying genealogy is a single typ branching process (which may be time non-homogeneous) and Assumption 2(a) can be easily checked. Moreover $m_n := m_n(x, \mathbf{e}, \mathcal{X}) = \Pi_{j=0}^{n-1} m(x, T^j\mathbf{e})$ does not depend on x. The law of large numbers of the Theorems given in the two previous section then rely (only, up to some moment assumptions) on the weak ergodicity of the auxiliary Markov chain Y_n whose kernel transition simplifies as

$$Q_n(x, \mathbf{e}, dy) := m_1(x, \mathbf{e}, dy) \frac{1}{m_1(x, \mathbf{e}, \mathcal{X})} = P(x, \mathbf{e}, dy).$$

Moreover $W_n = Z_n/m_n$ is (a.s. with respect to the environment) a martingale which converges to a positive finite limit on the non extinction event, since we have here L^2 assumptions. Then we obtain

$$\mathbb{P}_{\mathbf{e}}\left(\left|\frac{f_n.Z_n(f)}{Z_n(\mathcal{X})} - \mu_n(f)\right| \geq \eta \; ; \forall n \in \mathbb{N}, Z_n(\mathcal{X}) > 0\right) \stackrel{n\to\infty}{\longrightarrow} 0.$$

We recover here classical weak law of large numbers for proportions of individuals with a given trait for Markov chains along trees, such as Galton-Watson trees [G07] (Section 2.2), [DM10] (Theorem 1.3) and branching processes in random environment [BH13] (Theorem 3.2).

3.4.2 Using the spectral gap of the mean operator (in fixed environment).

In this section, the environment is fixed, so we can set $\mathbb{P} := \mathbb{P}_{\mathbf{e}}$ and

$$m_n(x,.) := m_n(x,\mathbf{e},.), \qquad m_n(\mu,.) = \int_{\mathcal{X}} \mu(dx) m_n(x,.)$$

for any $x \in \mathcal{X}$ and $\mu \in \mathcal{M}(\mathcal{X})$. We consider a subspace X of $\mathcal{M}(\mathcal{X})$ stable by addition which contains $\mathcal{M}_1(\mathcal{X})$. We endow X with a norm $\| \, . \, \|_X$ and assume that there exists $c > 0$ such that $\| \mu \|_X \leq c\mu(\mathcal{X})$ for any $\mu \in X$ and that $\mu \to m_1(\mu,.)$ is a bounded endomorphism on $(X, \| \, . \, \|_X)$. We denote by X' the topological dual of X and we assume that $(\mu \to m_1(.,\mathcal{X})) \in X'$ and the following spectral properties.

Assumption 3. There exists $(\lambda, \mu_0, f_0) \in (1,\infty] \times \mathcal{M}_1(\mathcal{X}) \times X'$ such that $f_0(\mu) > 0$ for any non zero measure μ of X and

$$m_1(\mu_0,.) = \lambda \mu_0(.), \qquad f_0(m_1(.,dx)) = \lambda f_0(.).$$

Moreover, there exists $a < \lambda$ and $c > 0$ such that

$$\| m_n(\mu,.) - \lambda^n f_0(\mu)\mu_0(.) \|_X \leq c a^n \| \mu - f_0(\mu)\mu_0(.) \|_X \, .$$

When $\mathcal{X}$ is finite, $\mathcal{M}(\mathcal{X})$ and X' are identified to vectors and f_0 has positive coefficients. Then, Perron Frobenius theory ensures that the previous Assumptions hold if the matrix given by the mean operator m_1 is aperiodic and irreducible. We refer to [S01] for details and extension to a denumerable state space $\mathcal{X}$. Moreover Krein Rutman Theorem allows to tackle the non-denumerable framework when the mean operator is compact and positive. A usual case corresponds then to identify f_0 with a (positive) measurable function and $f_0(m_1(.,dx)) = m(.,f_0)$. Let us finally note that several technics in analysis allow to go beyond these assumptions via the decompositions of the operator, see [MS14], where an overview in of the results in the continuous time framework is given.
The previous assumption ensures that for any non-negative function f such that $\mu_0(f) \in (0,\infty)$ and any $\mu \in X$,

$$m_n(\mu,f) \sim \lambda^n f_0(\mu)\mu_0(f) \qquad \text{and} \qquad m_n(\mu,\mathcal{X}) \sim f_0(\mu)\lambda^n \qquad (n \to \infty).$$

Moreover $f_0(Z_n)/\lambda^n$ is a martingale, so it converges a.s. to a finite r.v. Checking that the limit is non-degenerated is delicate in general. Classical $N \log N$ moment assumptions for single type population [LPP95] can be extended to the case $\#\mathcal{X} < \infty$ [KLPP97] and to more general framework. We refer to [A00] for extension of the $N \log N$ moment assumption and to [M67] for L^2 asymptotic behavior of $Z_n(.)$ when $\mathcal{X}$ is denumerable and to [C11] for a counterpart in continuous time. Here, we can use the L^2 computations of the previous section to derive the following statement for a general state space $\mathcal{X}$. Similarly, one could derive a strong law of large number from Theorem 2.

Corollary 3. *Let $f \in \mathcal{B}(\mathcal{X})$ bounded and $x \in \mathcal{X}$. If Assumption 3 hold and $\sup_{y\in\mathcal{X}} \mathbb{E}(N(y)^2) < \infty$, then*
i) $Z_n(\mathcal{X})/m_n(x,\mathcal{X})$ is bounded in L^2;
ii) $f_0(Z_n)/\lambda^n$ converges a.s. to $W \in [0,\infty)$ and $\mathbb{P}(W > 0) > 0$;
iii) $Z_n(f)/Z_n(\mathcal{X}) \longrightarrow \mu_0(f)$ as $n \to \infty$, $\mathbb{P}_{\delta_x}$ a.s. on the event $\{W > 0\}$.

Proof. Using Assumption 3, we obtain (12). Recalling $\lambda > 1$, (13) is satisfied. Then we can apply Lemma 6 and (2) hold. It ensures that $Z_n(\mathcal{X})/m_n(x,\mathcal{X})$ is bounded in L^2 and so does $f_0(Z_n)/\lambda^n$ since $f_0 \in X'$ is bounded and $\| Z_n \|_X \leq cZ_n(\mathcal{X})$. We deduce that the martingale limit of $f_0(Z_n)/\lambda^n$ is non-degenerated and $i - ii)$ are proved. Let us now focus on

$$Q_n(x,f) = m_n(\delta_x, f)/m_n(\delta_x,\mathcal{X}).$$

Using the second part of Assumption 3 with f_0 bounded, there exist constant c', c'' such that for every $y \in \mathcal{X}$

$$|Q_n(y,f) - \mu_0(f)| \leq c'(a/\lambda)^n \| \delta_y - f_0(\delta_y)\mu_0 \|_X \leq c''(a/\lambda)^n$$

and

$$|Q_n(y,f) - Q_n(z,f)| \leq 2c''(a/\lambda)^n.$$

It ensures that condition (11) hold since $a < \lambda$ and we can apply Theorem 4 to get

$$\frac{Z_n(f) - Z_n(\mathcal{X})\mu_0(f)}{m_n(x,\mathcal{X})} \stackrel{n\to\infty}{\longrightarrow} 0 \qquad \mathbb{P}_{\delta_x} \quad \text{a.s.}$$

Adding that $\liminf_{n\to\infty} Z_n(\mathcal{X})/m_n(x,\mathcal{X}) > 0$ a.s. on the event $\{W > 0\}$ since f_0 is bounded ends up the proof. □

3.4.3 Geometric ergodicity (in potentially varying environment).

We need here an additional assumption on the transition semi-groups, namely we require a contraction property, which can be handled by classical technics relying on Doeblin and Lyapounov type assumptions, see in particular [MT09, HM08, M13]. We consider the space of probabilities $\mathcal{M}_1(\mathcal{X})$, which we endowed with a distance d bounded by 1 such that for every bounded measurable function f, there exists $C > 0$ such that $|\mu(f) - \nu(f)| \leq Cd(\mu,\nu)$.

Corollary 4. *Assume that there exist C, A_n such that for any $\mathbf{e} \in E$, $x, y \in \mathcal{X}$, $\lambda, \mu \in \mathcal{M}_1(\mathcal{X})$ and $n \geq 0$,*

$$m_n(y,\mathbf{e},\mathcal{X}) \leq C(\mathbf{e})m_n(x,\mathbf{e},\mathcal{X}), \qquad d(Q_n(\lambda,\mathbf{e},.), Q_n(\mu,\mathbf{e},.)) \leq A_n(\mathbf{e})d(\lambda,\mu) \tag{15}$$

and

$$\sum_{n\geq 1} \frac{1 \wedge D(x, T^{n-1}\mathbf{e})}{m_n(x,\mathbf{e},\mathcal{X})} < \infty, \qquad \sum_{n\geq 0}\prod_{k=0}^{n} A_k(T^{n-k}\mathbf{e})^2 < \infty.$$

Then, for each $\mathbf{e} \in E$, $Z_n(\mathcal{X})/m_n(x,\mathbf{e},\mathcal{X})$ *is bounded in* $L^2_{\mathbf{e},\delta_x}$ *and for every measurable bounded function* f,

$$\frac{Z_n(f) - \mu_n(f)Z_n(\mathcal{X})}{m_n(x,\mathbf{e},\mathcal{X})} \overset{n\to\infty}{\longrightarrow} 0 \qquad \mathbb{P}_{\mathbf{e},\delta_x} \ a.s.$$

Proof. The proof is an application of Lemma 6 and Theorem 4. Indeed, we first use Lemma 6 to check that (10) hold and $Z_n(\mathcal{X})/m_n(x,\mathbf{e},\mathcal{X})$ is bounded in $L^2_{\mathbf{e},\delta_x}$. Moreover, by induction we obtain

$$d(Q_{i,n}(\lambda,\mathbf{e},.) - Q_{i,n}(\mu,\mathbf{e},.)) \leq \Pi_{k=i}^{n-1} A_{n-k}(T^k\mathbf{e})d(\lambda,\mu) \leq \Pi_{k=i}^{n-1} A_{n-k}(T^k\mathbf{e}),$$

since d is bounded by 1. Then

$$|Q_{i,n}(\lambda,\mathbf{e},f) - Q_{i,n}(\mu,\mathbf{e},f)| \leq C\Pi_{k=i}^{n-1} A_{n-k}(T^k\mathbf{e}).$$

Adding that the right-hand-side is summable allows us to get (11) and apply Theorem 4 and conclude. □

Let us provide now some tractable sufficient conditions to apply this result.

If there exists $M : E \to [1,\infty)$ such that for all $x, y \in \mathcal{X}, \mathbf{e} \in E$,

$$m_1(x,\mathbf{e},.) \leq M(\mathbf{e})m_1(y,\mathbf{e},.),$$

then (15) hold and the corollary can be applied. Such an assumption is motivated in particular by reproduction-dispersion model on islands (i.e. compact set). Indeed, we obtain the first part of (15) using

$$\begin{aligned} m_n(x,\mathbf{e}) = \int_{\mathcal{X}} m_1(x,\mathbf{e},dz)m_{n-1}(z,T\mathbf{e}) &\leq M(\mathbf{e})\int_{\mathcal{X}} m_1(y,\mathbf{e},dz)m_{n-1}(z,T\mathbf{e}) \\ &\leq M(\mathbf{e})m_n(y,\mathbf{e}), \end{aligned}$$

whereas the second part of (15) comes from

$$\begin{aligned} Q_n(x,\mathbf{e},A) &= \int_{\mathcal{X}} \frac{m_1(x,\mathbf{e},dz)}{m_n(x,\mathbf{e})} m_{n-1}(z,T\mathbf{e},A) \\ &\leq M(\mathbf{e})^2 \int_{\mathcal{X}} \frac{m_1(y,\mathbf{e},dz)}{m_n(y,\mathbf{e})} m_{n-1}(z,T\mathbf{e},A) \leq M(\mathbf{e})^2 Q_n(y,\mathbf{e},A) \end{aligned}$$

and d can be chosen as the total variation distance $d(\lambda,\mu) = \sup_{\|f\|_\infty \leq 1} \big| \int_{\mathcal{X}} f(x)\lambda(dx) - \int_{\mathcal{X}} f(x)\mu(dx) \big|$.

As a more general framework and motivation for further works, we can rewrite the second part of (15) in terms of the kernel transition P of the original model. Indeed, we can use the first part of (15) and the branching property to write

$$C(\mathbf{e})^{-1} m(x,\mathbf{e})m_{n-1}(y,T\mathbf{e}) \leq m_n(x,\mathbf{e}) \leq C(\mathbf{e})m(x,\mathbf{e})m_{n-1}(y,T\mathbf{e}).$$

Recalling that $Q_n(x,\mathbf{e},dy) = m(x,\mathbf{e})P(x,\mathbf{e},dy)m_{n-1}(y,T\mathbf{e})/m_n(x,\mathbf{e})$, we get

$$C(\mathbf{e})^{-1} P(x,\mathbf{e},dy) \leq Q_n(x,\mathbf{e},dy) \leq C(\mathbf{e})P(x,\mathbf{e},dy),$$

so that the second part of (15) can be derived from control on the semigroup P.

3.5 Further comments and applications

We now mention some links with classical branching models and provide few other applications.

3.5.1 About multitype branching processes

When the state space $\mathcal{X}$ is finite, the process Z is a multitype branching process and much finer results can be obtained. In particular, the limit behavior of $Z_n/m_n(x, \mathbf{e}, \mathcal{X})$ is known, see e.g. [KLPP97] in fixed environment and [C89] in random environment. Let us yet stress that we provide in Lemma 2 a slightly different spine decomposition than [KLPP97], without projection with respect to the eigenvector associated to the mean operator. In particular, in a work in progress, we are using such an expression for controlling the distribution of sampled cells at a fixed times and deriving estimations of parameters for cell division. Finally, we note that the previous results (and in particular Corollary 4) can be applied to branching processes in varying environment, when for example the mean growth rate decreases to 1 to mimic the effect of resources limitation.

In the two next subsections, we focus on the neutral case, which means that the reproduction law does not depend on the trait (see also Section 3.4.1). Let us deal with the good renormalizing function f_n which allows to get law of large numbers.

3.5.2 About branching random walks and random environment

Branching random walks have been largely studied from the pioneering works of Biggins (see e.g. [B77]) and central limit theorems have been obtained to describe the repartition of the population for large times [B90].
For branching random walks (possibly in varying environment in time and space), the auxiliary Markov chain Y is a random walk (possibly in varying environment in time and space). To get law of large numbers for Z_n, one can then check that some convergence in law

$$(Y_n - a_n)/b_n \Rightarrow W$$

where the limit W does not depend on the initial state $x \in \mathcal{X}$. Then we can use Theorem 3 with $f_n(x) = (x - a_n)/b_n$ to obtain the asymptotic proportion of individuals whose trait x satisfies $f_n(x) \in [a, b]$. It is given by $\mathbb{P}(W \in [a, b])$ as soon as $\mathbb{P}(W \in \{a, b\}) = 0$. Thus it can be used when the auxiliary process satisfies a central limit theorem. We refer to [BH13] Section 3.4 for some examples in the case when the reproduction law does not depend on the trait $x \in \mathcal{X}$ and the environment is stationary ergodic in time.
One can actually directly derive some (rougher) law of large numbers directly from the speed of random walk (in environment), i.e. use $Y_n/a_n \Rightarrow v$. As as example, we recall that in dimension 1 the random walk in random environment Y may be sub-ballistic and $b_n = n^\gamma$ with $\gamma \in (0, 1)$.
We finally mention [N11] when the offspring distribution is chosen in an i.i.d. manner for each time n and location $x \in \mathbb{Z}$.

3.5.3 About Kimmel's cell infection model and non-ergodicity

In the Kimmel's branching model [B08] for cell division with parasite infection, the auxiliary Markov chain Y_n is a Galton-Waston in (stationary ergodic) random environment. For example, in the case when no extinction is possible, i.e. $\mathbb{P}_1(Y_1 > 0) = 1$, under the usual integrability assumption we have

$$Y_n/\Pi_{i=0}^{n-1} m_i \stackrel{n\to\infty}{\longrightarrow} W \in (0, \infty) \qquad \text{a.s.}$$

where m_i is the (random) mean number of offsprings in a cell line in generation i. We note that the distribution of W depends on the initial value of Y. But

$$\log(Y_n)/n \stackrel{n\to\infty}{\longrightarrow} \mathbb{E}(\log m_0) \qquad \text{a.s.}$$

and the limit here does not depend on Y_0 any longer. So we get the ergodic property required to use use Theorem 3 with $f_n(x) = \log(x)/n$. We obtain that the proportion of cells in generation n whose number of parasites is between $\exp([\mathbb{E}(\log m_0) - \epsilon]n)$ and $\exp([\mathbb{E}(\log m_0) + \epsilon]n)$ goes to 1 in probability, for every $\epsilon > 0$. This yields some first new result on the infection propagation, which could be improved by additional work.
Soon as the number of parasites in a cell can be equal to zero, i.e. $\mathbb{P}_1(Y_1 = 0) > 0$, ergodicity is failing and some additional work is needed. Using monotonicity argument, one may still conclude, see [B08] for an example.

4 Local densities and extremal particles.

We deal now with local densities and the associated ancestral lineages. More precisely, we focus on the number of individuals whose trait belongs to some set A_n in generation n, when $n \to \infty$.
We have proved the many-to-one formula

$$\mathbb{E}(Z_n(A_n)) = m_n(x, \mathbf{e}, \mathcal{X}) Q_{0,n}(x, \mathbf{e}, A_n)$$

in the previous section. We have then checked that the ergodicity of $Q_{0,n}$ ensures that for fixed $A \in \mathcal{B}_{\mathcal{X}}$, $Z_n(A)/Z_n(\mathcal{X}) - Q_{0,n}(x, \mathbf{e}, A) \to 0$ under some conditions.
Now we wish to compare the asymptotic behaviors of

$$Z_n(A_n) \quad \text{and } m_n(x, \mathbf{e}, \mathcal{X}) Q_{0,n}(x, \mathbf{e}, A_n), \quad \text{when } Q_{0,n}(x, \mathbf{e}, A_n) \to 0.$$

In particular, we are studying the links between the local densities $Z_n(A_n)$ for large times and the large deviations events of $Q_{0,n}$, i.e. with the asymptotic behavior of $Q_{0,n}(x, \mathbf{e}, A_n)$.
Such questions have been well studied for branching random walks from the pioneering work of Biggins [B77]. We we refer to [R93, HS09] for related results and to [S08] for reviews on the topic. We also mention [CP07b, N12] for the random environment framework.
The upper bound for such results comes directly from Markov inequality and we are working on the lower bound. As usual, we could then derive the rough asymptotic behavior of the extremal (minimal or maximal position) individual. It covers classical (rough) results for branching random walks on the speed of the extremal individual (at the log scale). We provide some other examples motivated by cell's infection model, where the associated deviation strategy is more subtle. We finally mention that $Z_n(A_n)$ may be negligible compared to $m_n(x, \mathbf{e}, \mathcal{X}) Q_{0,n}(x, \mathbf{e}, A_n)$.

4.1 A general lower-bound

Let us derive a counterpart of Lemma 4 to get an a.s. lower bound for the number of individuals in generation n whose trait belong to A_n. We need here a more subtle construction than the subpopulation constructed by keeping the individuals whose trait is in some A_{n_k} for suitable generations n_k. This way works e.g. for branching random walks (see the pioneering works of Biggins) but the applications below are providing different pictures.
Let $x \in \mathcal{X}, \mathbf{e} \in E$, $(t_i : i \geq 0)$ be an increasing sequence of integers, $D_i \subset \{t_i, t_{i+1}, \ldots\}$ and $A_i, B_i \subset \mathcal{X}$. We set the event

$$\mathcal{T} := \left\{ \forall i \geq 0 : Z_{t_i}(B_i) > 0;\ \liminf_{i\to\infty} \frac{Z_{t_{i+1}}(B_{i+1})}{Z_{t_i}(B_i)} > 1 \right\}.$$

We assume now that $\mathbb{P}(\mathcal{T}) > 0$ so the following statement makes sense.

Lemma 7. *Assume that there exist measures $(\nu_i : i \geq 0)$ and ν and positive integers $(m_{i,n} : i, n \geq 0)$such that $\int_0^\infty l\nu(dl) < \infty$ and for all $i \geq 0$, $x \in B_i$, $l \geq 0$, $n \in D_i$,*

$$\mathbb{P}_{\delta_x, T^{t_i}\mathbf{e}}\left(\frac{Z_{n-t_i}(A_n)}{m_{i,n}} \geq l\right) \geq \nu_i[l,\infty], \qquad \nu_i[l,\infty) \leq \nu[l,\infty). \tag{16}$$

Then

$$\liminf_{i\to\infty} \inf_{n \in D_i} \frac{Z_n(A_n)}{Z_{t_i}(B_i) m_{i,n} \int_0^\infty l\nu_i(dl)} \geq 1$$

$\mathbb{P}_{\delta_x,\mathbf{e}}$ *a.s. on the event $\mathcal{T}$.*

In the applications we have in mind (see below), $\liminf_{i\to\infty} \int_0^\infty l\nu_i(dl) > 0$, $\cup_i D_i = \mathbb{N}$, $m_{i,n} \sim m_{n-t_i}(x_i, T^i\mathbf{e}, A_n)$ for some $x_i \in B_i$ and

$$Z_n(A_n) \approx Z_{t_i}(B_n) m_{t_i, n-t_i} \approx m_n(x, \mathbf{e}, A_n)$$

where $\approx$ is meaning an equivalence at a logarithm scale.

Proof. We use that for $i \geq 0$ and $n \in D_i$, the branching property and (16) ensure that

$$Z_n(A_n) \geq m_{i,n} \sum_{k=1}^{Z_{t_i}(B_i)} X_{i,k},$$

where $(X_{i,k} : k \geq 0)$ are i.i.d. r.v. distributed as ν_i. Then,

$$Z_n(A_n) \geq N_i m_{i,n} \frac{1}{N_i} \sum_{k=1}^{N_i} X_{i,k}$$

with $N_i := Z_{t_i}(B_i)$. Thanks to the domination of ν_i by ν, we can use Lemma 4 to get

$$\frac{1}{N_i} \sum_{k=1}^{N_i} X_{i,k} - \int_0^\infty l\nu_i(dl) \overset{i\to\infty}{\longrightarrow} 0 \quad \text{a.s.}$$

on the event $\mathcal{T}$. It ends up the proof. $\square$

4.2 Logarithm estimates of local densities

Let us derive the asymptotic behavior of the number of individuals in generation n whose trait belong to A_n. For safe of simplicity of the statement and regarding the applications below, we focus on the logarithm scale.

Proposition 1. *Let A_n be a sequence of subsets of $\mathcal{X}$. Then,*

$$(i) \qquad \limsup_{n\to\infty} \frac{1}{n} \log Z_n(A_n) \leq \limsup_{n\to\infty} \frac{1}{n} \log m_n(x, \mathbf{e}, A_n) \quad \mathbb{P}_{\mathbf{e},\delta_x} \ a.s.$$

(ii) We assume that there exist $t_i \geq 0$, $m_n > 0$, a non-zero measure with finite first moment ν and $i(n) \in \mathbb{N}$ such that for every $n \in \mathbb{N}$, $t_{i(n)} \leq n$ and

$$\mathbb{P}_{\delta_{b_{i(n)}}, T^{i(n)}\mathbf{e}} \left(\frac{Z_{n-t_{i(n)}}(A_n)}{m_n} \geq l \right) \geq \nu[l, \infty]. \tag{17}$$

Then

$$\liminf_{n\to\infty} \frac{1}{n} \log Z_n(A_n) \geq \limsup_{n\to\infty} \frac{1}{n} \log m_n$$

$\mathbb{P}_{\mathbf{e},\delta_x}$ *a.s. on the event* $\mathcal{T} = \big\{\forall i \geq 0 : Z_{t_i}(B_i) > 0, \ \liminf_{i\to\infty} Z_{t_{i+1}}(B_{i+1})/Z_{t_i}(B_i) > 1\big\}$.

Proof. As for branching random walks, the upper bound comes directly from Markov inequality. Indeed, for all $\varrho, \eta > 0$,

$$\mathbb{P}_{\mathbf{e},\delta_x}(Z_n(A_n) \geq \exp((\varrho+\eta)n)) \quad \leq \quad \exp(-(\varrho+\eta)n) m_n(x, \mathbf{e}, A_n).$$

So letting $\varrho = \limsup_{n\to\infty} \frac{1}{n} \log m_n(x, \mathbf{e}, A_n)$ ensures that the right hand side is summable. Then Borel-Cantelli lemma yields the a.s. upper bound (i).

The second part comes by applying Lemma 7 with $D_i := \{n : i(n) = i\}$, $m_{i,n} := m_n$ and $\nu_i := \nu$. Taking the logarithm and using $\int l\nu(dl) > 0$ with $\cup_i D_i = \mathbb{N}$ yields the results $\square$

This result ensures that at the logarithm scale, $Z_n(A_n)$ behaves as its mean. In particular, it allows to derive the speed of the extremal individuals in the population, following the standard arguments used e.g. for Branching random walks.
To explain how to check the assumptions required in this Theorem, let us provide L^2 sufficient conditions, in the same vein as Section 3.3. For safe of simplicity and applications, the results are given for monotone models.

4.3 L^2 computations, monotone branching Markov chains and large deviations

By now we assume that $\mathcal{X}$ is ordered by $\leq$ and with a slight abuse, we are denoting $[b, \infty] := \{x \in \mathcal{X} : x \geq b\}$.

Definition 2 (Monotonicity)**.** *We say that the branching Markov chain Z is monotone if for all $x \leq y$, $\mathbf{e} \in E$ and $a \in \mathcal{X}$, we have*

$$\mathbb{P}_{\delta_x,\mathbf{e}}(Z_1([a,\infty)) \geq l) \leq \mathbb{P}_{\delta_y,\mathbf{e}}(Z_1([a,\infty)) \geq l) \qquad (l \geq 0).$$

We consider the event

$$\mathcal{T} := \left\{ \forall i \geq 0 : Z_{t_i}([b_i, \infty]) > 0,\ \liminf_{i\to\infty} \frac{Z_{t_{i+1}}([b_{i+1}, \infty])}{Z_{t_i}([b_i, \infty])} > 1 \right\}$$

for some fixed increasing sequence $(t_i : i \geq 0)$ and elements b_i of $\mathcal{X}$.

We define the following measure for $k, n \geq 0$ and $a, b \in \mathcal{X}$,

$$\mu(k, n, a, b) := \mathbb{P}_{\delta_a, T^k\mathbf{e}}(Z_{n-k}([b, \infty)) \in .)$$

which counts the number of individuals with trait larger than b in generation n, which come from one single individual with trait a in generation k. For some measure μ, we note $\bar{\mu}$ its mean and $\hat{\mu}$ its normalized variance. Thus,

$$\bar{\mu}(k, n, a, b) = m_{n-k}(a, T^k\mathbf{e}, [b, \infty)), \qquad \hat{\mu}(k, n, a, b) = \mathbb{E}_{\delta_a, T^k\mathbf{e}}\left(\left[\frac{Z_{n-k}([b, \infty))}{\bar{\mu}(k, n, a, b)} \right]^2 - 1 \right).$$

Theorem 5. *Assume that Z is monotone and there exists $\varrho > 0$ such that*

$$\lim_{n\to\infty} \frac{1}{n} \log m_n(x, \mathbf{e}, [a_n, \infty)) = \varrho.$$

In addition, we assume that

$$(i) \qquad \liminf_{i\to\infty} \bar{\mu}(t_i, t_{i+1}, b_i, b_{i+1}) > 1, \quad \sum_{i\geq 0} \frac{\hat{\mu}(t_i, t_{i+1}, b_i, b_{i+1})}{\prod_{j=0}^{i-1} \bar{\mu}(t_j, t_{j+1}, b_j, b_{j+1})} < \infty.$$

(ii) for every $\epsilon > 0$, there exists $i(n) \in \mathbb{N}$ going to ∞, $\psi(n) \in \mathbb{N}$, increasing sequences $(t_{n,j})_{j\geq 0}$, $b_{n,j} \in \mathcal{X}$ such that $t_{n,\psi(n)} = n$, $b_{n,\psi(n)} = b_n$,

$$\liminf_{n\to\infty} \frac{1}{n} \sum_{j<\psi(n)} \log \bar{\mu}(t_{n,j}, t_{n,j+1}, b_{n,j}, b_{n,j+1}) \geq \varrho - \epsilon$$

and

$$\sup_n \sum_{j<\psi(n)} \frac{\hat{\mu}(t_{n,j}, t_{n,j+1}, b_{n,j}, b_{n,j+1})}{\prod_{l=0}^{j-1} \bar{\mu}(t_{n,l}, t_{n,l+1}, b_{n,l}, b_{n,l+1})} < \infty.$$

Then $\mathbb{P}(\mathcal{T}) > 0$ and

$$\lim_{n\to\infty} \frac{1}{n} \log Z_n(A_n) = \varrho \qquad \mathbb{P}_{\mathbf{e},\delta_x} \quad \textit{a.s. on the event } \mathcal{T}.$$

The proof relies on a coupling with two branching processes in varying environment. We note that the environment of the branching process may be varying even if that of orginial process does not. The finiteness of the series in (i) ensures that the associated martingale is non-degenerate, so that $\mathbb{P}(\mathcal{T}) > 0$. We recall the classical $L \log L$ criterion for the non-degenerescence of the martingale limit in the homogeneous case [LPP95]. Here, we use L^2 moment conditions in varying environment and refer to [G76] for more general (but less tractable) conditions.
For the applications, we shall use $\mathbb{E}(Z_{t_{i+1}-t_i}(B_{i+1})^2) \leq \mathbb{E}(Z_{t_{i+1}-t_i}(\mathcal{X})^2)$ and the estimates of Lemma 3 to check that the series above is finite.

Proof. We recall that a branching process in varying environment is an extension of Galton-Watson process X where the reproduction law μ_i in generation i may depend on i. Then, assuming that $\bar{\mu}_i \in (0,\infty)$,

$$\frac{X_n}{\Pi_{i<n}\bar{\mu}_i}$$

is a martingale and by orthogonality

$$\mathbb{E}\left(\left[\frac{X_n}{\mathbb{E}(X_n)}\right]^2 - 1\right) \leq \sum_{i<n} \frac{\hat{\mu}_i}{\prod_{j=0}^{i-1}\bar{\mu}_j}. \tag{18}$$

We use now a coupling argument by considering a subpopulation whose trait in generation t_i is larger than b_i. It is obtained simply by deleting the individuals (and their descendants) whose trait in generation i is not larger than b_i. Using the monotonicity assumption, the number of individuals Y_i that remain in generation i satisfies

$$Y_{i+1} \geq \sum_{j=1}^{Y_i} N_{i,j} \quad \text{a.s.}$$

for some r.v. $(N_{i,j} : j \geq 1)$ which can be chosen i.i.d. with common distribution $\mu(t_i, t_{i+1}, b_i, b_{i+1})$ and independent of $(N_{k,j} : k < i, j \geq 1)$. Then, Y is a.s. larger than a branching process in varying environment with reproduction law $\mu(t_i, t_{i+1}, b_i, b_{i+1})$ and the martingale

$$\frac{Y_n}{\Pi_{i\leq n}\mu(t_i, t_{i+1}, b_i, b_{i+1})}$$

converges a.s. to a non-negative r.v W. This martingale is bounded in L^2 due to Assumption (i) and (18). Then $\mathbb{E}(W) = 1$ and $\mathbb{P}(\mathcal{T}) > 0$ since

$$\liminf_{i\to\infty} \frac{Z_{t_{i+1}}([b_{i+1},\infty])}{Z_{t_i}([b_i,\infty])} \geq \liminf_{i\to\infty} \mu(t_i, t_{i+1}, b_i, b_{i+1}) > 1$$

a.s. on the event $\{W > 0\}$. Now we use similarly that

$$\text{under } \mathbb{P}_{\delta_{b_{i(n)}}, T^{i(n)}\mathbf{e}}, \qquad Z_{n-t_{i(n)}}([a_n,\infty]) \geq R^n_{\psi(n)} \qquad \text{a.s.,}$$

where $(R^n_l : l = 0, \ldots, \psi(n))$ is a branching process in varying environment started with one individual and successive reproduction laws $\mu(t_{n,l}, t_{n,l+1}, b_{n,l}, b_{n,l+1})$. Combining (18) and the second part of (ii), we obtain that

$$W_n := \frac{R^n_{\psi(n)}}{\prod_{j<\psi(n)} \bar{\mu}(t_{n,j}, t_{n,j+1}, b_{n,j}, b_{n,j+1})}$$

is bounded L^2. Adding that $\mathbb{E}(W_n) = 1$, Paley-Zygmund inequality ensures that there exists a non-zero measure ν such that

$$\mathbb{P}(W_n \geq l) \geq \nu[l,\infty).$$

Then

$$\mathbb{P}_{\delta_{b_{i(n)}}, T^{i(n)}\mathbf{e}}\left(\frac{Z_{n-t_{i(n)}}([a_n,\infty))}{\prod_{j<\psi(n)} \bar{\mu}(t_{n,j}, t_{n,j+1}, b_{n,j}, b_{n,j+1})} \geq l\right) \geq \nu[l,\infty)$$

and we can apply Proposition 1 $i - ii$). With our assumptions, it ensures that $\limsup_{n\to\infty} \log(Z_n(A_n))/n \geq \varrho$ a.s. and $\liminf_{n\to\infty} \log(Z_n(A_n))/n \leq \varrho - \epsilon$ a.s. on the event $\mathcal{T}$ for each $\epsilon > 0$. Then $\mathcal{T} \subset \{\lim_{n\to\infty} \frac{1}{n} \log Z_n(A_n) = \varrho\}$ and the proof is complete. $\square$

In the applications below, we consider the neutral case with fixed environment and

$$m_n(x, \mathbf{e}, A) = m^n Q_{0,n}(x, \mathbf{e}, A) = m^n P^n(x, A),$$

where P^n is the nth convolution of the kernel P introduced in Section 2. To check the conditions of the previous Theorem, we consider $t_i = ip$ and a subdivision $(t_{n,j} : j = 1, \ldots, \psi(n))$ of $[i(n), n]$ with step p and require that the second moment of the reproduction law is bounded ($\sigma := \sigma(\mathbf{e}, x) < \infty$). Then the conditions for $(i) - (ii)$ are satisfied as soon as

$$\liminf_{i\to\infty} m^p P^p(b_i, [b_{i+1}, \infty)) > 1, \qquad \liminf_{i,n\to\infty} m^p P^p(b_{n,i}, [b_{n,i+1}, \infty)) > 1$$

The applications which follow show how to find $(b_i, b_{i,n})$ for some relevant $A_n = [a_n, \infty]$ or $A_n = [-\infty, a_n]$. It underlines the natural link between the local density $Z_n([a_n, \infty])$ and the trajectory $(b_{n,i} : i \leq \psi(n))$ associated to the large deviation event $\{Y_n^{(n)} \geq a_n\}$.

4.4 Motivations and applications

We first give some details on a motivating example for which straight line and non straight line may appear for the large deviation strategy $(b_{n,i} : i \leq \psi(n))$. We then mention other challenging questions.

4.4.1 Kimmel's branching model

We refer to [B08] for a complete description of the model and the motivations. The population of individuals is a binary tree of cells and the trait is the number of parasites of the cell. The auxiliary Markov process Y is then a branching process in random environment. Monotonicity (see Definition 2) is a direct consequence of the branching property of Y. Tackling the local densities by means of the previous Theorem (only) requires to control the event $\{Y_n \in [a_n, \infty]\}$.
A first motivating question in [B08] is to count the number of infected cells in the subcritical case, which means that Y is a.s. absorbed in finite time. Three regimes appear in this case [GKV03] and in the weak subcritical case

$$\mathbb{P}(Y_n > 0) \sim cn^{-3/2}\gamma^n,$$

where $\gamma < \mathbb{E}_1(Y_1)$. Let us denote by $N_n^* = Z_n(\{1, 2, \ldots\})$ the number of infected cells in generation n. The mean number of infected cells $\mathbb{E}(N_n^*)$ is equal to $2^n \mathbb{P}(Y_n > 0)$ and obtaining a.s. results on the asymptotic behavior of N_n* was left open is this regime. Theorem 5 ensures that if $2\gamma > 0$, the number N_n^* of infected cells in generation n satisfies

$$\frac{1}{n} \log(N_n^*) \stackrel{n\to\infty}{\longrightarrow} \log(2\gamma) \qquad \text{a.s.}$$

on the event when the whole population of parasites survives. Indeed, this Theorem is applied for p large enough such that $2^p \mathbb{P}_1(Y_p > 0) > 1$ and $\log \mathbb{P}_1(Y_p > 0) \geq p(\log(\gamma) - \epsilon)$,

$A_n = [1, \infty]$, $b_i = 1, b_{n,l} = 1$, $i(n) = o(n)$ chosen such that .
Second, when counting the number of cells infected less than the typical cell in the supercritical regime, the problem is now linked to the lower large deviation of branching processes in random environment Y_n, i.e. to

$$\mathbb{P}(1 \leq Y_n \leq \exp(n\theta)), \qquad \text{where} \;\; \theta < \mathbb{E}(\log m(\mathcal{E}))$$

and the way this large deviation event is realized. We refer to [BB12] for detailed results. Here again Theorem 5 allows to determine the a.s. behavior of the number of cells whose number of parasites is between 1 and $\exp(\theta n)$. It is worth noting that for this question the associated trajectory is not a straight line and $b_{n,l}$ indeed depend on n. It is given by a continuous function which is piecewise affine.

4.4.2 Comments on branching random walks and random environment

We can recover here the classical convergence

$$\frac{1}{n} \log Z_n[an, \infty) \overset{n\to\infty}{\longrightarrow} \log(m) - \Lambda(a)$$

where $a \geq \mathbb{E}(X)$, $\log(m) > \Lambda(a)$ and Λ is the rate function associated to the random walk $S = \sum_{i=0}^{n-1} X_i$, see e.g. [S08].
One can extend this result to offsprings distribution in time varying environment and random walks in varying environment using the last Corollary and large deviations of random walks in varying environment. Here $b_i = aip$, $b_{n,l} = alp + b_{i(n)}$, $i(n) = o(n)$. We refer in particular to [Z04] for results on quenched and annealed large deviations of random walk in random environment.

4.4.3 Perspectives and Extremal individuals

A main motivation for the last results is the control of local densities in cell division models for aging [G07, DM10], for damages [ES07] or infection. In particular, we are now considering with Valère Biteski Penda the local densities for bifurcating autoregressive processes [G07]. More precisely, we are deriving the number of cells whose growth rate is becoming large when the generation is going to ∞.
The results given should also help to handle the non neutral framework. More precisely, the trait space is the age of the cell (and maybe some additional aging factor, such as the number of ERCs), which influences the mortality of the cell.
Finally, let us recall that determining the asymptotic behavior of $Z_n([a_n, \infty))$ at the logarithm scale allows to derive the speed of the maximal trait in generation n among the population. Indeed, roughly speaking if $a_n(x)$ satisfies the assumptions of Theorem 5 with some rate $\rho(x)$:

$$\lim_{n\to\infty} \frac{1}{n} \log m_n(x, \mathbf{e}, [a_n(x), \infty)) = \varrho(x).$$

Then, for every x such that $\rho(x) > \log m$,

$$\limsup_{n\to\infty} \frac{\max\{X(u) : |u| = n\}}{a_n(x)} \leq 1 \qquad \mathbb{P}_{\delta_x, \mathbf{e}} \text{ a.s.}$$

and for every x such that $\rho(x) < \log m$,

$$\liminf_{n\to\infty} \frac{\max\{X(u) : |u| = n\}}{a_n(x)} \geq 1 \qquad \mathbb{P}_{\delta_x,\mathbf{e}} \text{ a.s.}$$

on some event whose probability is positive.

Acknowledgement. This work was partially funded by Chair Modelisation Mathematique et Biodiversite VEOLIA-Ecole Polytechnique-MNHN-F.X., the professorial chair Jean Marjoulet, the project MANEGE 'Modèles Aléatoires en Écologie, Génétique et Évolution' 09-BLAN-0215 of ANR (French national research agency).

References

[AS10] E. Aïdékon, Z. Shi (2010). Weak convergence for the minimal position in a branching random walk: a simple proof. *Periodica Mathematica Hungarica* 61 (2010) 43-54.

[AK98a] K. Athreya, H.J. Kang (1998). Some limit theorems for positive recurrent branching Markov chains I. *Adv. Appl. Prob.* 30(3). 693-710.

[AK98b] K. Athreya, H.J. Kang (1998). Some limit theorems for positive recurrent branching Markov chains II. *Adv. Appl. Prob.* 30(3). 711-722.

[AK71] K. B. Athreya, S. Karlin (1971). On branching processes with random environments II : Limit theorems. *Ann. Math. Stat.* 42. 1843-1858.

[A00] K. B. Athreya (2000). Change of measure of Markov chains and the $L \log L$ theorem for branching processes. *Bernoulli* Vol. 6, No 2, 323-338.

[B08] V. Bansaye (2008). Proliferating parasites in dividing cells : Kimmel's branching model revisited. *Annals of Applied Probability, Vol 18, Number 3, 967-996.*

[BB12] V. Bansaye, C. Boeinghoff (2012). Lower large deviations for supercritical branching processes in random environment. To appear in the *Proc. of Steklov Institute of Mathematics.*

[BDMT11] V. Bansaye, J.-F. Delmas, L. Marsalle and V.C. Tran (2011). Limit theorems for Markov processes indexed by continuous time Galton-Watson trees. *Ann. Appl. Probab.* Vol. 21, No. 6, 2263-2314.

[BL12] V. Bansaye, A. Lambert (2012). New approaches of source-sink metapopulations decoupling the roles of demography and dispersal. To appear in *Theor. Pop. Biology*

[BH13] V. Bansaye, C. Huang (2013). Weak law of large numbers for some Markov chains along non homogeneous genealogies. *Preprint* avialable via Arxiv.

[B77] J.D. Biggins (1977). Martingale convergence in the branching random walk. *J. Appl. Probab.* **14**, 25-37.

[B90] J. D. Biggins (1990). The central limit theorem for the supercritical branching random walk and related results. *Stoch. Proc. Appl.* **34**, 255-274.

[C11] B. Cloez (2011). Limit theorems for some branching measure-valued processes. *Avialable via* http://arxiv.org/abs/1106.0660.

[CRW91] B. Chauvin, A. Rouault, A. Wakolbinger (1991). Growing conditioned trees. *Stochastic Processes and their Applications* **39**, 117–130.

[C89] H. Cohn (1989). On the growth of the supercritical multitype branching processes in random environment. *Ann. Probab.* Vol. 17, No 3. 1118-1123.

[CP07a] F. Comets, S. Popov (2007). Shape and local growth for multidimensional branching random walks in random environment. *ALEA* **3**, 273-299.

[CP07b] F. Comets, S. Popov (2007). Shape and local growth for multidimensional branching random walks in random environment. *ALEA* **3**, 273-299.

[CY11] F. Comets, N. Yoshida (2011). Branching random walks in space-time random environment: survival probability, global and local growth rates. *Journal of Theor. Probab.*

[DZ98] A. Dembo, O. Zeitouni (1998). *Large deviations techniques and applications.* Applications of Mathematics (New York) 38 (Second edition ed.)

[DMS05] A. Dembo, P. Mörters, S. Sheffield (2005). Large deviations of Markov chains indexed by random trees. *Ann. Inst. H. Poincaré Probab. Statist.* 41, no. 6, 971?996.

[DM10] J.-F. Delmas, L. Marsalle (2010). Detection of cellular aging in a Galton-Watson process. *Stochastic Processes and their Applications* **120**, 2495–2519.

[E07] J. Engländer (2007). Branching diffusions, superdiffusions and random media. *Prob. Surveys.* **4**, 303-364.

[EK86] S.N. Ethier, T.G. Kurtz (1986). *Markov Processus, Characterization and Convergence.* John Wiley & Sons, New York.

[ES07] S.N. Evans, D. Steinsaltz (2007). Damage segregation at fissioning may increase growth rates: A superprocess model. *Theoretical Population Biology* **71**, 473-490.

[F71] W. Feller (1971). *An introduction to probability theory and its applications*, volume 1 and 2. Wiley.

[FK60] H. Furstenberg, H. Kesten (1960). Products of random matrices. *The Annals of Mathematical Statistics*, 31(2):457-469.

[GM05] N. Gantert, S. Müller .The critical Branching Markov Chain is transient. *Arxiv* http://arxiv.org/abs/math/0510556v1.

[GMPV10] N. Gantert, S. Müller, S. Popov, M. Vachkovskaia (2010). Survival of branching random walks in random environment. *Journal of Theor. Probab.*, **23**, 1002-1014.

[GKV03] J. Geiger, G. Kersting, V. A. Vatutin (2003). Limit theorems for subcritical branching processes in random environment. *Ann. Inst. Henri Poincaré (B).* **39**, pp. 593–620.

[G99] J. Geiger (1999). Elementary new proofs of classical limit theorems for Galton-Watson processes. *J. Appl. Prob. 36*, 301-309.

[GB03] H.O. Georgii, E. Baake (2003). Supercritical multitype branching processes: the ancestral types of typical individuals. *Adv. in Appl. Probab.*, Vol. 35, No 4, 1090-1110.

[G76] R. T. Goettge (1976). Limit theorems for the supercritical Galton-Watson process in varying environments. *Math. Biosci.* 28, no. 1-2, 171-190.

[GRW92] L. G. Gorostiza, S. Roelly, A. Wakolbinger (1992). Persistence of critical multitype particle and measure branching processes. *Probability Theory and Related Fields*, Vol. 92, No 3, 313-335.

[G07] J. Guyon. Limit theorems for bifurcating Markov chains. Application to the detection of cellular aging. *Ann. Appl. Probab.* **17**, 1538–1569 (2007).

[HM08] M. Hairer, J. C. Mattingly (2008). Yet another look at Harris' ergodic theorem for Markov chains. *Avialable via* arXiv:0810.2777.

[HR12] S. C. Harris, M. I. Roberts (2012). The many-to-few lemma and multiple spines. *Avialable via* http://arxiv.org/abs/1106.4761.

[HR13] S. C. Harris, M. I. Roberts (2013). A strong law of large numbers for branching processes: almost sure spine events *Avialable via* http://arxiv.org/abs/1302.7199.

[HL11] C. Huang, Q. Liu (2011). Branching random walk with a random environment in time. *Preprint.*

[HS09] Y. Hu, Z. Shi (2009). Minimal position and critical martingale convergence in branching random walks, and directed polymers on disordered trees. *Ann. Probab.* 37 742-781.

[JN96] P. Jagers, O. Nerman (1996). The asymptotic composition of supercritical multi-type branch- ing populations. In *Séminaire de Probabilités*, XXX, volume 1626 of Lecture Notes in Math., pages 40-54. Springer, Berlin, 1996.

[K74] N. Kaplan (1974). Some Results about Multidimensional Branching Processes with Random Environments. *Ann. Probab.*, Vol. 2, No. 3., 441–455.

[K72] T. Kurtz (1972). Inequalities for law of large numbers. *Ann.Math.Statist.* 43,1874-1883.

[KLPP97] T. Kurtz, R. Lyons, R. Pemantle, Y. Peres (1997). A conceptual proof of the Kesten-Stigum theorem for multi-type branching processes. In *Classical and Modern Branching Processes*, ed. K. B. Athreya and P. Jagers. Springer, New York. 181-185.

[LPP95] R. Lyons, R. Pemantle, Y. Peres (1995). Conceptual proofs of $L \log L$ criteria for mean behavior of branching processes. *Ann. Probab.*, Vol. 23, No 3, 1125-1138.

[MT09] S. Meyn L. Tweedie (2009). *Markov Chains and Stochastic Stability*. Broché.

[M67] S.-T.C. Moy (1967). Extensions of a limit theorem of Everett Ulam and Harrison multi-type branching processes to a branching process with countably many types. *Ann. Math. Statist.* 38. 992-999.

[MS14] S. Mischler, J. Scher (2013). Spectral analysis of semigroups and growth-fragmentation equations. *Avialable on Arxiv* http://arxiv.org/abs/1310.7773.

[M13] F. Mukhamedov (2013). Weak ergodicity of nonhomogeneous Markov chains on noncommutative L1-spaces. *Banach J. Math. Anal.* Vol. 7, No. 2, 53 -73.

[N11] M. Nakashima (2011). Almost sure central limit theorem for branching random walks in random environment. *Ann. Appl. Probab.*, **21**(1), 351-373.

[N12] M. Nakashima (2013). Minimal Position of Branching Random Walks in Random Environment. *J. Theor Probab.* 26:1181-1217

[NJ84] O. Nerman, P. Jagers (1984). The stable double infinite pedigree process of supercritical branching populations. *Z. Wahrsch. Verw. Gebiete* 65 , no. 3, 445-460.

[R93] A. Rouault (1993). Precise estimates of presence probabilities in the branching random walk. *Stoch. Process. Appl.* 44, no. 1, 27-39.

[S01] E. Seneta (2006). *Non-negative matrices and Markov chains.* Springer Series in Statistics. No. 21

[S94] T. Seppäläinen (1994). Large deviations for Markov chains with Random Transitions. *Ann. Prob.* **22** (2), 713-748.

[S08] Z. Shi (2008). *Random walks and trees.* Lecture notes, Guanajuato, Mexico, November 3-7.

[T88] D. Tanny (1988). A necessary and sufficient condition for a branching process in a random environment to grow like the product of its means. *Stoch. Process. Appl.* 28, no. 1, 123-139.

[Y08] N. Yoshida (2008). Central limit theorem for random walk in random environment. *Ann. Appl. Probab.* **18** (4), 1619-1635.

[Z04] O. Zeitouni (2004). Random walks in random environment. *XXXI Summer school in probability, St Flour* (2001). Lecture notes in Math. 1837 (Springer) (2004), pp. 193?312.